

BAB II

TINJAUAN PUSTAKA

2.1 Data Mining

Data Mining adalah proses mengekstraksi informasi dan pola yang berharga dari sejumlah besar data (Halima & Benaboud, 2025). Berdasarkan pendapat dari (Wijaya, 2023), data mining adalah suatu tahapan pengolahan data berskala besar yang bertujuan untuk mengidentifikasi informasi serta struktur tersembunyi dari himpunan data berukuran besar melalui pendekatan seperti machine learning dan metode lainnya. Metode ini sering disebut sebagai *Knowledge Discovery in Databases* (KDD) atau dalam bahasa Indonesia dikenal dengan sebagai “Penemuan Pengetahuan”, karena kemampuannya dalam mengungkap pola tersembunyi dari data yang dianalisis. Tujuan Utama data mining yaitu mengeksplorasi data sebagai pengetahuan yang bermanfaat menjadi dasar untuk meningkatkan kualitas pengambilan keputusan.

Data mining banyak diterapkan di berbagai bidang, dimulai dari keuangan, kesehatan, hingga pemasaran dan pemerintahan. Salah satu penerapannya adalah dalam analisis data penerima bantuan sosial, seperti Bantuan Langsung Tunai, untuk mengidentifikasi kelompok masyarakat yang memenuhi syarat menerima bantuan berdasarkan data yang tersedia, seperti informasi demografis, status pekerjaan, pendapatan, dan kondisi sosial-ekonomi lainnya.

Penentuan algoritma yang digunakan dalam data mining perlu mempertimbangkan tujuan analitis yang hendak dicapai serta relevansinya terhadap masing-masing fase dalam tahapan *Knowledge Discovery in Databases*

(*KDD*) (Bachtiar & Mahradianur, 2023). Proses ini memungkinkan mengidentifikasi pola, korelasi, serta kecenderungan yang tidak tampak secara langsung dan yang tidak mudah diidentifikasi dengan analisis manual (Provinsi et al., 2023). Menurut (Putra, 2020), Data mining adalah proses yang umum digunakan dalam berbagai bidang, seperti ekonomi, sosial dan kesehatan, untuk mengolah data yang belum terstruktur menjadi informasi yang berguna dalam mendukung keputusan strategis.

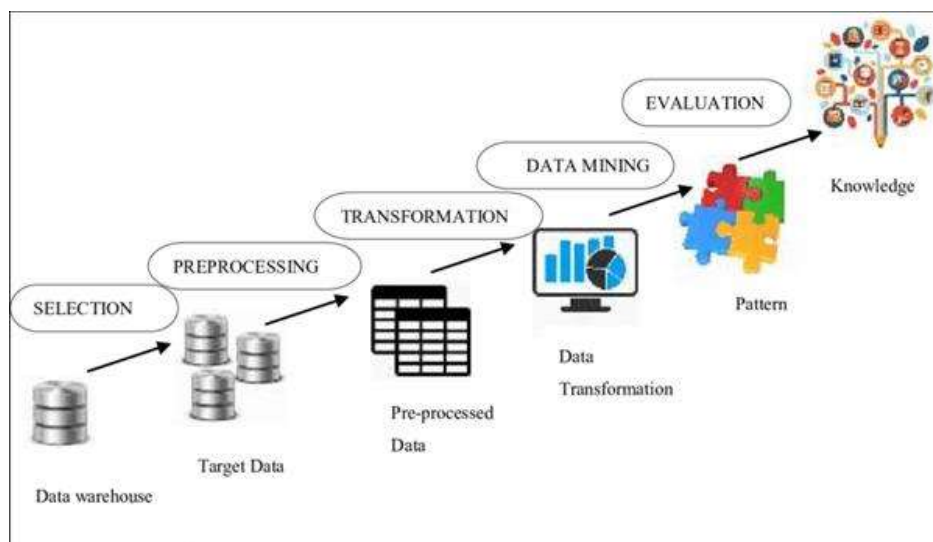
Proses perolehan data, pembersihan data, pemilihan fitur, penerapan algoritma dan evaluasi hasil adalah beberapa langkah penting dalam proses data mining. Pengumpulan data dari beberapa sumber adalah langkah pertama dalam proses, setelah itu dibersihkan untuk menghilangkan informasi yang tidak akurat atau tidak diperlukan. Selanjutnya, Ciri-ciri yang paling relevan kemudian dipilih untuk pemeriksaan tambahan. Selanjutnya, metode data mining untuk mengekstraksi keterkaitan atau pola tersembunyi di dalam data.

Teknik-teknik yang umum digunakan pada pemrosesan data lanjutan yang mencakup pengelompokan berdasarkan kategori (klasifikasi), regresi, pengelompokan tanpa label (klasterisasi), dan pencarian pola hubungan (asosiasi). Dalam metode klasifikasi, data dimasukkan ke dalam kelompok yang telah ditetapkan sebelumnya, sementara regresi digunakan untuk memprediksi nilai kontinu.

Secara garis besar, data mining tidak hanya digunakan untuk menganalisis data, tetapi juga untuk memahami dan memanfaatkan pengetahuan yang tersimpan dalam kumpulan data besar. Oleh karena itu, proses pengambilan

keputusan yang lebih optimal dapat dilakukan pada berbagai sektor. Hal ini memungkinkan pengguna melihat hubungan yang mungkin tersembunyi di antara banyak informasi, yang mempermudah calon penerima dalam membuat keputusan yang lebih baik.

Dalam penelitian ini, data mining digunakan untuk mengklasifikasikan kelayakan masyarakat sebagai penerima BLT. Metode Naive Bayes digunakan untuk melakukan klasifikasi ini yang memungkinkan pengelompokan data berdasarkan kriteria tertentu, seperti pendapatan dan status pekerjaan.



Sumber : <https://sis.binus.ac.id/2021/09/30/proses-data-mining-kdd/>

Gambar 2. 1 Metodologi Data Mining

2.1.1 Tahapan Proses Data Mining

A. Selection (Pemilihan Data)

Tahap pertama dalam proses data mining adalah Selection, yang melibatkan pemilihan data yang relevan dari berbagai sumber atau gudang data. Pada tahap ini, data dipilih untuk dianalisis berdasarkan tujuan penelitian atau masalah yang

ingin diselesaikan. Data yang dipilih ini harus mencakup informasi yang relevan dan penting untuk tahap-tahap selanjutnya.

B. Preprocessing (Prakiraan Data)

Setelah data dipilih, tahap selanjutnya adalah preprocessing atau penyaringan data, dimana informasi tersebut dibersihkan dari nilai tidak tersedia, salinan ganda, maupun kesalahan lainnya. Tidak relevan atau tidak konsisten data dibuang, yang menghasilkan data yang lebih akurat dan dapat diandalkan. Ini memungkinkan pengguna untuk melihat hubungan yang mungkin tersembunyi di antara banyak data, yang memungkinkan mereka dalam mengambil keputusan yang lebih optimal.

Penelitian ini menggunakan data mining untuk mengkategorikan kelayakan masyarakat sebagai penerima BLT. Metode Naive Bayes digunakan untuk melakukan klasifikasi ini, yang memungkinkan pengelompokan data berdasarkan kriteria tertentu, seperti pendapatan dan status pekerjaan.

C. Transformation (Transformasi Data)

Setelah melalui proses pembersihan, data kemudian masuk ke tahap transformasi. Di tahap ini, data diolah kembali agar lebih siap digunakan dalam proses analisis. Proses ini dapat mencakup penyesuaian skala nilai (normalisasi), penggabungan beberapa sumber data, maupun peringkasan informasi melalui agregasi. Tahap ini sangat penting untuk menyiapkan data dalam bentuk yang ideal untuk dianalisis lebih lanjut.

D. Data Mining (Penambangan Data)

Setelah data dipilih, dibersihkan, dan mengalami proses transformasi, tahap berikutnya adalah melakukan data mining. Pada tahap ini, metode seperti klasifikasi, klusterisasi, asosiasi, maupun regresi diaplikasikan guna menganalisis data lebih dalam. Tujuan utamanya adalah menemukan pola atau keterkaitan yang tidak terlihat sebelumnya, yang bisa dimanfaatkan sebagai landasan untuk membuat keputusan yang lebih tepat.

E. Evaluation (Evaluasi)

Setelah pola atau informasi ditemukan, tahap Evaluation dilakukan untuk menilai hasil yang telah diperoleh dari proses data mining. Pada bagian ini, hasil analisis dievaluasi agar dapat dipastikan bahwa pola yang diperoleh relevan dan konsisten. Evaluasi ini penting untuk menentukan apakah model atau hasil analisis benar-benar berguna untuk tujuan yang telah ditetapkan. Hasil evaluasi akan menunjukkan apakah model perlu disesuaikan atau diperbaiki.

F. Knowledge (Pengetahuan)

Tahapan akhir dalam proses data mining adalah tahap *Knowledge*, yang berfokus pada pemanfaatan hasil analisis untuk mendukung pengambilan keputusan atau memperdalam pemahaman terhadap suatu permasalahan. Pada tahap ini, informasi yang berhasil digali selama proses data mining dimanfaatkan secara praktis, seperti menentukan kelompok masyarakat yang memenuhi syarat untuk menerima bantuan sosial, atau menyusun kebijakan baru berdasarkan pola-pola yang telah ditemukan.

2.1.2 Fungsi Data Mining

Dalam berbagai industri, seperti pemerintahan, bisnis, kesehatan, dan banyak lagi data mining telah menjadi alat yang sangat penting. Beberapa jurnal mengidentifikasi beberapa fungsi utama data mining yang relevan untuk klasifikasi kelayakan penerima bantuan sosial (BLT).

a. Klasifikasi

Menurut (Azlil Huriah & Dienwati Nuris, 2023) Dalam data mining dan machine learning, klasifikasi adalah salah satu elemen penting. Proses ini merupakan teknik pengelompokan data, di mana setiap informasi yang dianalisis sudah memiliki label atau kelas tujuan tertentu. Data dikategorikan berdasarkan kelompok yang sudah didefinisikan sebelumnya selama proses klasifikasi. Model yang digunakan dalam proses klasifikasi membantu memetakan informasi ke dalam kategori tertentu berdasarkan variabel yang tersedia, seperti kelayakan masyarakat untuk menerima bantuan BLT. Pada algoritma seperti *Naïve Bayes* sangat efektif dalam menangani masalah klasifikasi berbasis data yang besar dan kompleks.

b. Prediksi

Prediksi merupakan proses memperkirakan nilai-nilai baru berdasarkan data historis, dengan tujuan mengurangi perbedaan antara data aktual dan nilai yang diperkirakan oleh model (Windy Mardiyah et al., 2024). Fungsi ini menjadi sangat penting ketika analisis difokuskan untuk memperkirakan siapa saja yang akan menerima bantuan di masa mendatang berdasarkan karakteristik yang dimiliki. Dalam konteks Bantuan Langsung Tunai (BLT), prediksi digunakan

untuk mengidentifikasi calon penerima berdasarkan kriteria yang telah ditetapkan sebelumnya. Dengan penerapan algoritma data mining, proses prediksi ini dapat meningkatkan efisiensi dan ketepatan dalam penyaluran bantuan sosial karena mempertimbangkan berbagai variabel secara lebih sistematis dan berbasis data.

c. Asosiasi dan Penemuan Pola

Menurut (Erwansyah et al., 2021) Asosiasi digunakan untuk mengidentifikasi pola perilaku dari kejadian-kejadian tertentu atau mengidentifikasi keterkaitan yang terjadi antara berbagai peristiwa pada suatu proses. Fungsi asosiasi adalah menggali pola atau keterkaitan yang tidak tampak secara langsung antar variabel pada dataset. Algoritma Apriori menunjukkan bahwa proses eksplorasi data mampu mengungkap aturan-aturan asosiasi yang menghubungkan variabel dalam dataset yang lebih besar. Dalam konteks BLT, asosiasi ini bisa mengidentifikasi hubungan antara atribut seperti usia, pendapatan, status pekerjaan, dan kemungkinan seseorang untuk menerima bantuan. Misalnya, ada kemungkinan bahwa individu yang berpendapatan rendah dan tidak bekerja lebih mungkin untuk memenuhi kriteria penerima bantuan BLT.

d. *Clustering* (Pengelompokan)

Clustering adalah teknik data mining yang digunakan untuk memisahkan data menjadi berbagai kategori berdasarkan kesamaan karakteristik di berbagai kumpulan berdasarkan hal-hal yang sebanding (Hendrastuty, 2024). Metode ini tanpa menggunakan label kelas yang telah ditentukan sebelumnya menemukan struktur dalam data. Dalam konteks penerima bantuan BLT, teknik clustering

dapat digunakan untuk mengelompokkan individu berdasarkan karakteristik serupa dan memberikan bantuan secara lebih terarah.

Clustering dapat sangat berguna dalam menganalisis data demografis untuk menemukan pola yang tidak terlihat langsung dalam distribusi bantuan. Dengan kata lain, clustering memungkinkan pemerintah untuk mengelompokkan wilayah atau kelompok masyarakat yang membutuhkan bantuan sosial berdasarkan ciri-ciri yang serupa, seperti tingkat pendapatan atau jumlah tanggungan keluarga.

e. Deteksi Anomali

Menemukan data dengan perilaku yang berbeda dari kebiasaan dikenal sebagai deteksi anomali. Dalam data mining, teknik ini berguna untuk mengidentifikasi data yang tidak sejalan dengan pola umum dan memperlihatkan kemungkinan terjadinya kesalahan atau tindakan penipuan (Saputra et al., 2024).

Pada penelitian BLT, deteksi anomali dapat digunakan untuk menemukan penerima yang mungkin melakukan kecurangan atau data yang tidak sesuai, seperti pengajuan bantuan dari individu yang tidak memenuhi kriteria yang seharusnya.

f. Evaluasi Model

Dalam data mining, proses evaluasi sangat penting karena memastikan bahwa model yang digunakan benar-benar menghasilkan output yang akurat dan bermanfaat. Dari studi mengenai evaluasi model klasifikasi, disarankan penggunaan metrik yang tepat misalnya tingkat akurasi, presisi, recall, serta F1-score. Tujuan penggunaan metrik tersebut adalah untuk menjamin bahwa sistem

yang dibangun mampu memberikan performa yang stabil serta memiliki kemampuan dalam menghasilkan hasil yang akurat.

2.1.3 Klasifikasi (*Classification*)

Klasifikasi merupakan metode dalam data mining yang berfungsi untuk menentukan atau memprediksi kategori entitas menggunakan data dengan label sebelumnya. Proses ini dilakukan dengan membangun model prediktif menggunakan data latih, yaitu data yang telah memiliki kelas tertentu. Selanjutnya, data baru yang belum diketahui kelasnya diklasifikasikan menggunakan model yang telah dilatih (Kegiatan, 2022).

Menurut (Yendra et al., 2020), Dengan tujuan utama dalam menentukan kategori entitas yang tidak dikenali, maka klasifikasi adalah tahapan membangun fungsi atau model yang mampu membedakan kelas satu dengan kelas lainnya. Proses ini sangat bermanfaat dalam banyak bidang, seperti dalam menentukan apakah seseorang layak menerima Bantuan Langsung Tunai (BLT), karena membantu mengidentifikasi secara tepat individu atau keluarga yang berhak menerima bantuan berdasarkan sejumlah variabel yang dianalisis.

Beberapa algoritma klasifikasi yang sering digunakan meliputi *Naïve Bayes*, *Decision Tree*, dan *Support Vector Machine* (SVM). Pada konteks penyaluran BLT, penerapan metode-metode ini memungkinkan proses analisis yang lebih terstruktur, seperti dalam mengevaluasi variabel seperti pendapatan, jumlah tanggungan keluarga, tingkat pendidikan, hingga status pekerjaan. Algoritma klasifikasi Naive Bayes dipilih untuk penelitian ini karena merupakan metode berbasis perhitungan peluang yang dikenal sederhana namun cukup efektif. Naïve

Bayes didasarkan pada gagasan bahwa setiap atribut dalam data saling independen (bebas).

Meskipun asumsi ini tidak selalu dipenuhi dalam implementasi nyata, metode ini masih dapat menghasilkan keluaran yang cukup akurat, terutama untuk data berukuran besar dengan struktur yang tidak terlalu kompleks.

Keunggulan utama Naïve Bayes terletak pada kecepatan dan efisiensinya dalam membangun model prediktif. Algoritma ini tidak memerlukan proses pelatihan yang rumit dan dapat langsung digunakan untuk mengklasifikasikan data baru setelah model selesai dihitung. Kondisi ini menjadikan metode tersebut sangat sesuai untuk digunakan dalam proses seleksi penerima Bantuan Langsung Tunai (BLT), terutama karena informasi yang diterapkan dapat berasal dari banyak sumber dan perlu diolah dengan baik.

Pada penentuan kelayakan BLT, Naïve Bayes digunakan untuk menghitung probabilitas kelayakan berdasarkan sejumlah atribut seperti penghasilan bulanan, pekerjaan, jumlah tanggungan, status kepemilikan rumah, dan tingkat pendidikan. Model akan mengestimasi kemungkinan seseorang masuk ke dalam kategori “Layak” atau “Tidak Layak” berdasarkan nilai-nilai atribut tersebut. Kategori dengan probabilitas tertinggi kemudian dijadikan keputusan akhir sistem.

Selain itu, algoritma Naïve Bayes juga memiliki keunggulan dalam menangani permasalahan ketidakseimbangan data, karena pendekatannya didasarkan pada perhitungan probabilitas dan tidak terpengaruh oleh distribusi jumlah masing-masing kelas. Metode ini juga cukup tahan terhadap

ketidakkonsistenan data, sehingga dapat memberikan hasil prediksi yang stabil meskipun digunakan pada data yang berubah-ubah.

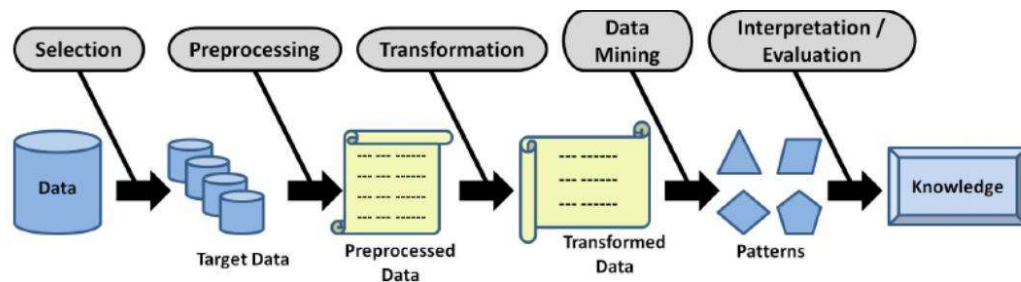
Dengan mempertimbangkan kemudahan implementasi, kecepatan pemrosesan, serta hasil klasifikasi yang cukup akurat, metode Naïve Bayes menjadi pilihan yang tepat dalam penelitian ini sebagai dasar pengambilan keputusan untuk menentukan kelayakan masyarakat sebagai penerima BLT secara otomatis dan objektif.

Hasilnya, distribusi bantuan menjadi lebih tepat sasaran, mengurangi kesalahan, dan memastikan bahwa bantuan sampai ke yang paling membutuhkan. Keunggulan dari Metode klasifikasi dapat menghasilkan hasil yang terstruktur dengan baik dan mudah dipahami. Model dengan dihasilkan bersifat transparan dan dapat dijelaskan kepada pihak-pihak terkait dalam proses pengambilan keputusan. Selain itu, klasifikasi juga mampu menangani dataset yang tidak seimbang (*imbalanced data*) dengan pendekatan seperti oversampling (menambahkan data pada kelas minoritas) atau undersampling (Data dikurangi untuk kelas mayoritas).

Jika dibandingkan dengan metode lain dalam data mining seperti clustering, association, prediction, dan deteksi anomali, klasifikasi memiliki pendekatan yang lebih langsung terhadap masalah penentuan kategori. Clustering, seperti yang dijelaskan oleh Hendrastuty (2024), bertujuan untuk mengelompokkan data tanpa label berdasarkan kesamaan karakteristik. Karena clustering bersifat unsupervised, maka metode ini kurang sesuai untuk penentuan langsung penerima bantuan yang memerlukan label “Layak” atau “Tidak Layak”.

Sementara itu, association rules lebih fokus pada pencarian pola hubungan antar variabel dalam satu set data, seperti dalam transaksi belanja, dan tidak digunakan untuk klasifikasi kelayakan individu (- & Tamba, 2022). Prediction merupakan proses memperkirakan nilai atau hasil di masa depan, dan meskipun klasifikasi termasuk di dalamnya, tidak semua metode prediksi berbasis klasifikasi. Deteksi anomali digunakan untuk mengidentifikasi data yang menyimpang dari pola umum, yang lebih relevan dalam konteks keamanan atau kontrol kualitas daripada seleksi penerima bantuan. Dengan mempertimbangkan berbagai metode tersebut, klasifikasi menjadi pilihan yang paling sesuai untuk aplikasi BLT karena mampu menghasilkan keputusan yang relevan, akurat, dan dapat diandalkan. Dalam penelitian ini, metode klasifikasi digunakan untuk menentukan kelayakan masyarakat sebagai penerima Bantuan Langsung Tunai berdasarkan atribut-atribut seperti penghasilan, pekerjaan, jumlah tanggungan, status kepemilikan rumah, dan faktor relevan lainnya. Sistem klasifikasi selanjutnya akan mengelompokkan individu ke dalam kategori "Layak" atau "Tidak Layak" sebagai penerima bantuan, sehingga dapat mendukung pemerintah dalam proses penyaluran bantuan yang lebih adil dan tepat sasaran. Oleh sebab itu, penggunaan metode klasifikasi tidak hanya memberikan keuntungan secara teknis, selain itu, mendukung terciptanya pelayanan publik yang lebih adil dan efisien. Berdasarkan alasan tersebut, penelitian ini mengadopsi pendekatan klasifikasi dalam membangun sistem penentuan kelayakan penerima Bantuan Langsung Tunai (BLT), agar proses distribusi bantuan dapat dilakukan secara objektif, tepat sasaran, dan berkelanjutan.

2.2 Knowledge Discovery in Database (KDD)



Gambar 2. 2 Siklus Proses KDD

Sumber : <https://www.researchgate.net/figure/kd>

Dalam penelitian ini, metode klasifikasi digunakan dalam *Knowledge Discovery in Database* (KDD) merupakan pendekatan yang mengumpulkan pengetahuan melalui kumpulan informasi atau database yang sudah tersedia. Metode ini digunakan untuk menentukan status kelayakan masyarakat sebagai penerima Bantuan Langsung Tunai. Setelah itu, pengetahuan ini dapat digunakan untuk membuat keputusan. Proses penemuan pengetahuan dalam database ini dilakukan melalui beberapa tahapan yang saling berurutan dan harus dijalankan secara sistematis (Filki, 2022).

Knowledge Discovery in Database (KDD) mencakup berbagai aktivitas dalam lingkungan kerja, mulai dari tahap pengumpulan data hingga proses interpretasi hasil analisis. Dalam penelitian ini, Kelayakan penerima Bantuan Langsung Tunai (BLT) ditentukan oleh KDD melalui penerapan metode klasifikasi Naïve Bayes. Sistem berbasis web yang dikembangkan dirancang untuk mengimplementasikan tahapan KDD dalam mengolah data masyarakat, sehingga proses seleksi penerima bantuan dapat dilakukan secara otomatis, efisien, dan lebih objektif. Menurut (Nainel et al., 2020) Proses penemuan

pengetahuan dalam KDD (*Knowledge Discovery in Databases*) melibatkan akuisisi serta pemanfaatan data terdahulu guna mengidentifikasi pola, kecenderungan, maupun relasi dalam himpunan data berskala besar. Menurut (Pratama et al., 2021) Data mining, yang kerap diistilahkan sebagai proses ekstraksi pengetahuan dari basis data (KDD).

Menurut (Gunadi, 2022) *Knowledge Discovery in Databases* (KDD) didefinisikan sebagai salah satu proses untuk menelusuri dan menganalisis data skala besar untuk mengekstraksi informasi serta pengetahuan sebagai bernilai dengan bantuan teknologi komputer. KDD merupakan cabang ilmu yang berkaitan erat dengan basis data, dan dapat dijelaskan sebagai proses pengenalan pola yang menghasilkan pengetahuan yang dapat dipahami dari kumpulan data yang kompleks.

Proses ini terdiri dari beberapa tahapan yang memiliki tujuan untuk mengolah data awal menjadi informasi bernilai serta berguna bagi proses pengambilan keputusan. Tahapan utama dalam proses *Knowledge Discovery in Databases* (KDD) meliputi seleksi data, pra-pemrosesan, transformasi, analisis data, evaluasi serta pemaknaan pola, dan diakhiri dengan penyajian hasil pengetahuan. Dalam praktiknya, kegiatan pengolahan data pada KDD mencakup pembersihan data, integrasi data, pemilihan data yang relevan, hingga penyajian hasil pengetahuan. Saat ini, banyak sektor seperti bisnis, pendidikan, dan pemerintahan telah mengadopsi pemrosesan data secara terkomputerisasi. Data yang tersimpan dalam database memungkinkan sistem komputer untuk menjalankan proses tersebut secara efisien dan terstruktur.

2.3 Metode Naïve Bayes

Menurut (Nizam Fadli et al., 2021), pendekatan Naïve Bayes merupakan fondasi dari Teknik klasifikasi yang berasumsi bahwa setiap atribut atau variable bersifat saling bebas, suatu asumsi yang dikenal dengan sebutan “naïve”. Dalam penerapannya, Tingkat akurasi prediksi pada algoritma Naïve Bayes dapat meningkat apabila didukung oleh jumlah data pelatihan yang mencukupi. Kuantitas data pelatihan yang besar akan membantu model dalam menghasilkan klasifikasi yang lebih tepat. Pada konteks BLT, Naïve Bayes akan menghitung kemungkinan kelayakan penerima BLT berdasarkan atribut seperti pendapatan, jumlah tanggungan, dan status pekerjaan.

Dalam proses penentuan layak tidaknya masyarakat menerima Bantuan Langsung Tunai (BLT), pendekatan Naïve Bayes dimanfaatkan. Pendekatan ini bersandar pada prinsip probabilistik, dimana perhitungan peluang dilakukan terhadap masing-masing atribut pada data untuk secara otomatis menetapkan status kelayakan. Pendekatan ini dimaksudkan untuk menggantikan proses seleksi manual yang cenderung memakan waktu serta berpotensi mengandung unsur subjektivitas. Selain memiliki waktu pemrosesan yang cepat, metode ini juga memiliki struktur yang cukup sederhana sehingga mudah dipahami dan diimplementasikan. Penerapan algoritma naïve bayes pada platform web memungkinkan proses klasifikasi berjalan secara lebih efisien dan cepat. Naïve Bayes tergolong sebagai salah satu metode yang paling umum dimanfaatkan dalam aktivitas klasifikasi pada data mining.

Metode Naive Bayes menghitung kemungkinan setiap kelas yang mungkin (seperti, layak atau tidak layak) berdasarkan atribut yang ada dalam data input. Kelas yang memiliki nilai kemungkinan tertinggi dipilih untuk diklasifikasikan. Pendekatan ini sangat cocok diterapkan dalam sistem berbasis web karena waktu komputasi yang relatif singkat dan struktur algoritma yang sederhana, sehingga memungkinkan pengolahan data secara efisien dan real-time.

Tujuan utama dari metode *Naïve Bayes* adalah untuk menentukan probabilitas suatu peristiwa berdasarkan informasi probabilitas dari peristiwa lain yang sudah diketahui. Menurut (Mardiani et al., 2023), output klasifikasi dari sistem data mining akan memiliki nilai guna yang lebih tinggi apabila disampaikan secara informatif, menarik, dan mudah dimengerti oleh pihak yang menjadi target pemanfaatan data tersebut. Persamaan umum yang dimiliki pada algoritma klasifikasi *naïve bayes* menurut (Muhamad Adhi Mukti et al., 2023) dapat dilihat pada Persamaan dibawah ini.

Teorema Umum pada probabilitas Naïve Bayes :

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

Rumus 2. 1 Teorema Naive Bayes

Keterangan:

- a. **P (H | X)** : Probabilitas hipotesis **H** (misal: Layak) berdasarkan data **X**.
- b. **P (H)** : Probabilitas Awal kelas **H**.
- c. **P (X|H)** : Probabilitas data **X** muncul jika **H** benar.
- d. **P (X)** : Probabilitas Keseluruhan dari data **X**.

2.4 Klasifikasi Dengan Beberapa Metode

Salah satu pendekatan penting dalam pengolahan data untuk keperluan klasifikasi banyak diterapkan dalam berbagai jenis analisis, seperti melalui metode Naïve Bayes, Algoritma C4.5, maupun Support Vector Machine (SVM). Berikut ini disajikan perbandingan antara ketiga metode klasifikasi tersebut:

Menurut (Rayuwati et al., 2022), Teorema Naïve Bayes merupakan konsep dasar dalam statistika dan probabilitas yang pertama kali diperkenalkan oleh Thomas Bayes, seorang ilmuwan asal Inggris. Pendekatan Naïve Bayes termasuk dalam metode klasifikasi yang berlandaskan pada teorema tersebut. Teknik ini mengandalkan data historis untuk memperkirakan kemungkinan terjadinya suatu peristiwa di masa mendatang. Adanya asumsi kuat tentang independensi antar variabel atau atribut yang dianalisis adalah ciri utama algoritma Naive Bayes. Dalam penelitian ini, *Naïve Bayes* dipilih sebagai metode klasifikasi karena memiliki sejumlah keunggulan yang mendukung dalam proses penentuan kelayakan masyarakat penerima Bantuan Langsung Tunai (BLT) implementasi sistem dalam bentuk aplikasi web. *Naïve Bayes* merupakan algoritma sederhana dan efisien, memungkinkan proses data dengan cepat untuk memberikan hasil secara real-time. Algoritma ini juga mampu menangani data dengan banyak fitur, seperti pendapatan dan jumlah tanggungan, serta menghasilkan klasifikasi yang akurat meskipun data terbatas. Selain itu, *Naïve Bayes* cocok untuk klasifikasi biner, misalnya "layak" atau "tidak layak," sehingga hasilnya mudah diinterpretasi. Kemampuan algoritma ini menangani *missing values* serta resisten

terhadap *overfitting* membuatnya semakin ideal untuk sistem klasifikasi penerima BLT yang fleksibel dan responsif.

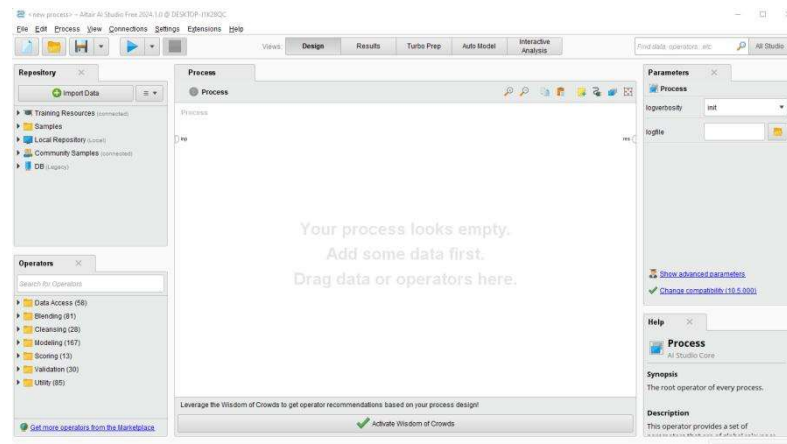
Menurut (Febriyani et al., 2021) Dalam banyak penelitian, algoritma C4.5 digunakan untuk merancang model model berbentuk decision tree. Metode ini termasuk ke dalam teknik klasifikasi dan prediksi yang sangat dikenal serta banyak digunakan. C4.5 bekerja dengan menyusun model klasifikasi dalam bentuk struktur pohon, di mana setiap data akan dipetakan ke dalam kelas tertentu berdasarkan nilai dari atribut-atribut yang dimilikinya. Dalam algoritma ini, pohon keputusan dibangun berdasarkan perhitungan *gain ratio*, yang memilih atribut dengan kontribusi informasi terbesar dalam membedakan kelas-kelas pada data.

Menurut (Octaviani et al., 2014) Metode klasifikasi Support Vector Machine (SVM) awalnya dibuat untuk menangani masalah klasifikasi yang dapat dipisahkan secara linier. Namun, seiring perkembangannya, SVM kini mampu menangani permasalahan non-linier melalui penerapan konsep *kernel*, yang memetakan data ke dalam ruang berdimensi tinggi. Dalam ruang tersebut, algoritma ini akan mencari sebuah *hyperplane* yang mampu memaksimalkan margin atau jarak antar kelas. SVM dikenal efektif dalam mengolah data berdimensi tinggi dan dapat digunakan baik untuk klasifikasi linier maupun non-linier dengan memanfaatkan fungsi kernel.

Dalam analisis klasifikasi kelayakan penerima Bantuan Langsung Tunai (BLT) yang dilakukan melalui situs web, metode Naive Bayes dianggap sesuai karena didukung oleh sejumlah variabel yang meningkatkan efisiensi dan

efektivitas proses klasifikasi tersebut. Pertama, *Naïve Bayes* menawarkan kesederhanaan dan kemudahan dalam interpretasi, sehingga memudahkan para pemangku kepentingan untuk memahami hasil klasifikasi tanpa memerlukan latar belakang teknis yang mendalam. Selain itu, metode ini sangat cepat dalam hal pelatihan dan prediksi, yang memungkinkan sistem untuk memberikan hasil secara real-time, suatu hal yang krusial dalam konteks pengambilan keputusan terkait bantuan sosial. Kinerja *Naïve Bayes* juga terbukti baik meskipun dengan data pelatihan yang terbatas, menjadikannya pilihan yang ideal ketika data yang tersedia tidak melimpah. Di samping itu, algoritma ini memiliki toleransi yang tinggi terhadap noise, yang sering kali terdapat dalam data sosial, sehingga meningkatkan akurasi klasifikasi. Berbeda dengan metode lain seperti SVM atau C4.5, yang mungkin lebih kompleks dan memerlukan waktu lebih lama untuk pelatihan serta tuning parameter, *Naïve Bayes* menawarkan solusi yang lebih efisien dan mudah diimplementasikan. Keunggulan lain dari *Naïve Bayes* adalah kemampuannya untuk memberikan output dalam bentuk probabilitas, yang memungkinkan pengambil keputusan untuk memahami tingkat kepastian dari klasifikasi yang dihasilkan. Dengan semua keunggulan ini, *Naïve Bayes* menjadi pilihan yang sangat relevan dan efektif untuk analisis klasifikasi kelayakan masyarakat penerima bantuan BLT.

2.5 RapidMiner



Gambar 2. 3 Program Rapidminer Studio

Sumber : <https://docs.rapidminer.com/9.9/studio/installation/>

Karena antarmuka yang mudah digunakan dan fitur drag-and-dropnya, Rapidminer telah menjadi perangkat populer yang kerap dimanfaatkan dalam kegiatan data mining maupun analisis data. Fitur-fitur ini membuatnya mudah digunakan bahkan oleh pengguna yang tidak terbiasa dengan pemrograman. Menurut (Sholeh et al., 2023), aplikasi ini mampu menangani berbagai tugas, seperti eksplorasi data, penggalian teks, serta analisis berbasis prediksi. Rapidminer memfasilitasi pembuatan model yang mendukung proses pengambilan keputusan melalui penerapan beragam teknik analisis, baik deskriptif maupun prediktif. Kelebihan lainnya adalah pengguna tidak perlu menulis kode secara manual, karena seluruh proses pembuatan model dapat dilakukan melalui operator-operator yang telah disediakan di dalam sistem. Pembuatan model dilakukan dengan menyusun dan menghubungkan berbagai operator yang relevan sehingga membentuk satu kesatuan proses. Penggunaan RapidMiner dalam pengembangan data mining telah banyak diterapkan oleh

sejumlah peneliti, dengan variasi datasheet, model, dan algoritma yang disesuaikan dengan kebutuhan masing-masing penelitian.

2.6 Metode Pengembangan Sistem

Menurut (Afiatul Ulumik & Irma Purnamasari, 2023), prototyping merupakan salah satu pendekatan dalam rekayasa perangkat lunak yang bertujuan untuk menggambarkan cara kerja sistem atau komponennya dalam lingkungan operasional sebelum memasuki fase pengembangan secara menyeluruh. Penelitian ini menggunakan metode prototyping sebagai pendekatan dalam pengembangan sistem. Metode ini dipilih karena bersifat iteratif, memungkinkan adanya interaksi yang aktif antara pengembang dan pengguna selama proses perancangan berlangsung. Pendekatan ini dinilai sesuai untuk sistem klasifikasi kelayakan masyarakat penerima Bantuan Langsung Tunai (BLT) berbasis web, karena memungkinkan adanya proses evaluasi dan perbaikan sistem secara berkelanjutan melalui masukan langsung dari pihak desa maupun operator pengelola data bantuan.

Metode prototyping memungkinkan pengembangan sistem yang lebih cepat dan fleksibel melalui pembuatan purwarupa (prototype) yang dapat langsung diuji oleh pengguna. Dengan demikian, sistem dapat disesuaikan secara berkala sampai benar-benar memenuhi kebutuhan fungsional yang diharapkan.

Proses prototyping yang digunakan dalam penelitian ini dapat dijelaskan sebagai berikut:

1. Tahapan dalam kebutuhan pengguna sistem dievaluasi. Ini termasuk prosedur input data masyarakat, pengolahan data menggunakan metode

Naive Bayes, dan penampilan hasil klasifikasi. Kebutuhan sistem diperoleh melalui studi literatur, observasi, serta wawancara dengan pihak-pihak terkait penyaluran bantuan, seperti aparat desa dan petugas dinas sosial.

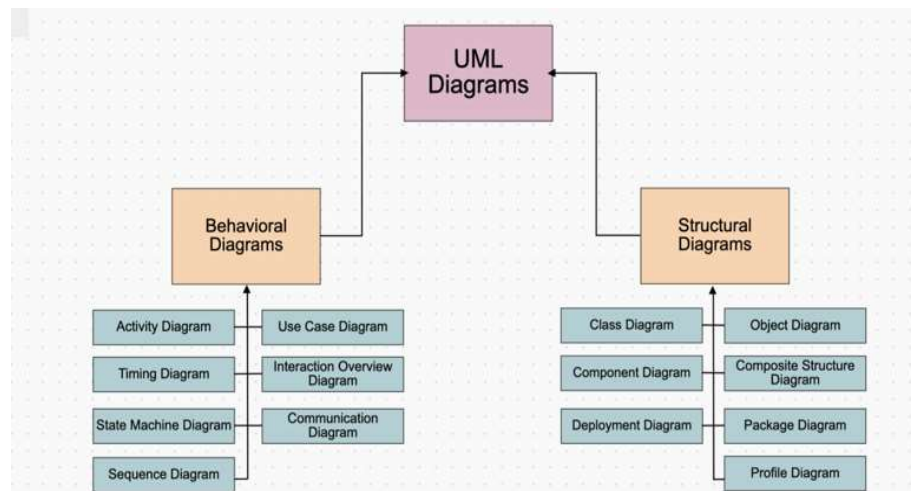
2. Perancangan Awal Sistem (Quick Design) Setelah kebutuhan dikumpulkan, dilakukan perancangan awal sistem. Desain ini meliputi tampilan antarmuka pengguna (*user interface*), struktur halaman web, dan alur sistem klasifikasi. Tujuannya adalah untuk memberikan gambaran awal tentang sistem yang akan dibuat.
3. Pembuatan Prototype Berdasarkan perancangan awal, dibuatlah prototype sistem berbasis web yang memiliki fungsi utama seperti form input data warga, pemrosesan klasifikasi menggunakan metode Naïve Bayes, dan tampilan hasil prediksi kelayakan berdasarkan jenis bantuan. Prototype ini merupakan cikal bakal sistem akhir, meskipun masih bersifat sederhana dan belum sepenuhnya optimal.
4. Evaluasi dan Umpan Balik Pengguna Prototype diuji langsung oleh pengguna sistem, yaitu pihak yang terlibat dalam proses seleksi penerima bantuan. Pengguna akan memberikan masukan terkait kemudahan penggunaan, kejelasan informasi, keakuratan klasifikasi, dan efisiensi alur kerja sistem. Masukan ini menjadi dasar untuk revisi dan pengembangan lebih lanjut.
5. Penyempurnaan Prototype Setelah mendapat umpan balik, prototype disempurnakan dari segi fitur, tampilan, maupun performa sistem.

Pengguna dapat melakukan proses ini berulang kali hingga calon penerima bantuan merasa sistem sesuai dengan kebutuhan mereka. Pada saat ini, juga dilakukan pemeriksaan tentang tingkat akurasi teknik pengklasifikasian data masyarakat menggunakan Naive Bayes.

6. Implementasi Sistem Final Setelah sistem dianggap layak dan sesuai, prototype diubah menjadi sistem akhir yang siap digunakan. Sistem ini kemudian dapat diimplementasikan secara penuh sebagai alat bantu pengambilan keputusan dalam menentukan kelayakan masyarakat sebagai penerima BLT secara cepat, akurat, dan berbasis web.

Metode prototyping mendukung pengembangan sistem BLT berbasis web yang tidak hanya cepat dan responsif, tetapi juga lebih sesuai dengan kondisi nyata di lapangan karena melibatkan pengguna dalam proses evaluasi secara langsung. Dengan demikian, sistem yang dibangun diharapkan mampu mempermudah proses seleksi penerima bantuan dan meningkatkan akurasi penyaluran dana bantuan kepada masyarakat yang benar-benar membutuhkan.

2.7 Unified Modeling Language (UML)



Gambar 2. 4 Diagram UML

Sumber : <https://itbox.id/blog/uml-adalah/>

Menurut (Abdillah, 2021), Unified Modeling Language (UML) merupakan bahasa pemodelan perangkat lunak yang telah mengalami standarisasi dan berfungsi sebagai acuan visual atau cetak biru dalam tahap perancangan sistem perangkat lunak. Serangkaian diagram yang diberikan oleh UML (Unified Modeling Language) digunakan untuk memvisualisasikan, merancang, dan mendokumentasikan berbagai aspek sistem, baik dari sisi struktur maupun perilaku. Penelitian ini difokuskan pada analisis kelayakan calon penerima Bantuan Langsung Tunai (BLT) melalui sistem yang diimplementasikan dalam bentuk aplikasi web. Proses kerja sistem, interaksi pengguna-sistem, dan struktur internal sistem digambarkan menggunakan UML. UML menyediakan cara visual untuk menggambarkan struktur dan perilaku sistem melalui berbagai jenis diagram.

Berdasarkan Penelitian (Fu'adi & Prianggono, 2022) menyatakan bahwa Unified Modeling Language (UML) memiliki total 13 jenis diagram yang dikelompokkan ke dalam tiga kategori utama. Berikut gambaran ringkas mengenai klasifikasi diagram tersebut.

a. Structure Diagram

Kategori ini mencakup sekumpulan diagram yang digunakan untuk merepresentasikan struktur statis dari sistem yang dimodelkan. Diagram yang termasuk dalam kelompok ini antara lain: *Class Diagram*, *Object Diagram*, *Component Diagram*, *Composite Structure Diagram*, *Package Diagram*, dan *Deployment Diagram*.

b. Behavior diagram

Diagram dalam kategori ini digunakan untuk memodelkan perilaku sistem atau menggambarkan perubahan yang terjadi dalam sistem seiring waktu. Diagram yang termasuk dalam kelompok ini meliputi: *Use Case Diagram*, *Activity Diagram*, dan *State Machine Diagram*.

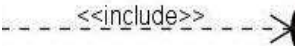
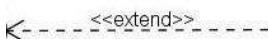
c. Interaction diagram

Kelompok diagram ini fokus pada penggambaran interaksi antara komponen-komponen sistem, baik antar subsistem maupun dengan sistem eksternal. Diagram yang termasuk dalam kategori ini adalah: *Sequence Diagram*, *Communication Diagram*, *Timing Diagram*, dan *Interaction Overview Diagram*.

2.7.1 Use case Diagram

UseCase diagram menunjukkan fungsi utama yang harus dimiliki oleh suatu sistem dan bagaimana pengguna (aktor) berinteraksi dengan sistem untuk mencapai tujuan tertentu. Dalam Use Case juga terdapat aktor, yang merupakan representasi dari entitas sistem atau user yang melakukan tugas di sistem (Ramdany, 2024).

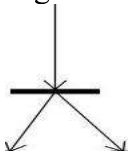
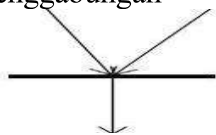
Table 2. 1 *Usecase Diagram*

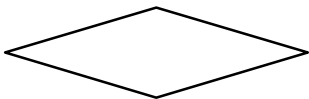
Simbol	Deskripsi
<i>Use case</i> 	menggambarkan fungsionalitas yang disediakan sistem sebagai unit-unit yang bertukar pesan antar unit dengan aktor, yang dinyatakan dengan menggunakan kata kerja
Aktor / <i>actor</i> 	<i>Actor</i> atau Aktor adalah <i>Abstraction</i> dari orang atau sistem yang lain yang mengaktifkan fungsi dari target sistem.
Asosiasi / <i>association</i> 	Asosiasi antara aktor dan <i>use case</i> , digambarkan dengan garis tanpa panah yang mengindikasikan siapa atau apa yang meminta interaksi secara langsung.
Asosiasi / <i>association</i> 	Asosiasi antara aktor dengan <i>use case</i> yang menggunakan panah terbuka untuk mengindikasikan bila aktor berinteraksi secara pasif dengan sistem.
<i>Include</i> 	<i>Include</i> , merupakan di dalam <i>use case</i> lain (<i>required</i>) atau pemanggilan <i>use case</i> oleh <i>use case</i> contohnya adalah pemanggilan sebuah fungsi program
<i>Extend</i> 	<i>Extend</i> , merupakan perluasan dari <i>use case</i> lain jika kondisi atau syarat terpenuhi

2.7.2 Activity Diagram

Activity Diagram dirancang untuk mempresentasikan urutan aktivitas atau proses kerja dalam sistem yang akan dioperasikan. Diagram ini berfungsi untuk memvisualisasikan alur proses serta mengorganisasi setiap tahap aktivitas yang berlangsung secara bertahap dan berurutan. Diagram ini terdiri dari berbagai elemen berbentuk khusus yang saling terhubung melalui panah, yang menunjukkan urutan aktivitas dari awal proses hingga selesai. (Simare Mare et al., 2022).

Table 2. 2 *Activity Diagram*

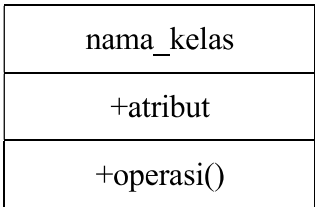
Simbol	Deskripsi
<i>Start Point</i> 	<i>Start Point</i> , diletakkan pada pojok kiri atas dan merupakan awal aktivitas
<i>End Point</i> 	<i>End Point</i> , akhir aktivitas
<i>Activities</i> 	<i>Activities</i> , menggambarkan suatu proses atau kegiatan bisnis
<i>Fork atau Percabangan</i> 	<i>Fork</i> atau percabangan, digunakan untuk menunjukan kegiatan yang dilakukan secara paralel untuk menggabungkan dua kegiatan paralel menjadi satu
<i>Join atau Penggabungan</i> 	<i>Join</i> (penggabungan) atau <i>rake</i> , digunakan untuk menunjukan adanya dekomposisi

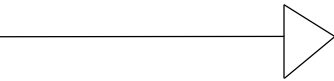
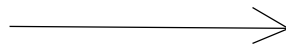
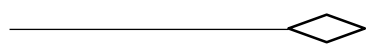
<i>Decision Points</i> 	<i>Decision points</i>, menggambarkan pilihan untuk pengambilan keputusan, <i>true</i> atau <i>false</i>
<i>Swimlane</i> 	<i>Swimlane</i>, pembagian <i>activity diagram</i> untuk menunjukan siapa melakukan apa.

2.7.3 Class Diagram

Jenis Class Diagram sangat penting dalam UML karena mampu memetakan sistem secara menyeluruh dengan memodelkan operasi, atribut, kelas, dan hubungan antar objek yang terlibat (Simare Mare et al., 2022).

Table 2. 3 Class Diagram

Simbol	Deskripsi
Kelas 	Kelas pada struktur sistem
Antar muka / <i>interface</i> 	Sama dengan konsep <i>interface</i> dalam pemrograman berorientasi objek
Asosiasi / <i>association</i> 	Relasi antarkelas dengan makna umum, asosiasi biasanya juga disertai dengan <i>multiplicity</i>
Asosiasi berarah / <i>directed association</i> 	Relasi antarkelas dengan makna kelas yang satu digunakan oleh kelas yang lain, asosiasi biasanya juga disertai dengan <i>multiplicity</i>

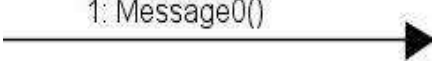
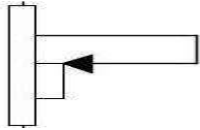
Generalisasi 	Relasi antarkelas dengan makna generalisasi-spesialisasi (umum khusus)
Kebergantungan / <i>dependency</i> 	Relasi antarkelas dengan makna kebergantungan antarkelas
Agregasi / <i>aggregation</i> 	Relasi antar kelas dengan makna semua – bagian (<i>whole-part</i>)

2.7.4 Sequence Diagram

Salah satu jenis diagram yang digunakan untuk menunjukkan alur proses yang dijalankan oleh pengguna adalah sequence diagram dalam sistem informasi, berdasarkan urutan waktu dari setiap tahapan aktivitas yang terjadi. Sequence Diagram menunjukkan bagaimana objek berinteraksi, disusun berdasarkan urutan waktu. Sequence Diagram menunjukkan urutan langkah yang diperlukan dalam suatu sistem. Hal ini dilakukan untuk mewujudkan skenario yang dijelaskan dalam Usecase Diagram (Amelia Sari Lubis et al., 2023).

Table 2. 4 Sequence Diagram

Simbol	Deskripsi
 Entity Class	<i>Entity Class</i> , merupakan bagian dari sistem yang berisi kumpulan kelas berupa entitas-entitas yang membentuk gambaran awal sistem dan menjadi landasan untuk menyusun basis data
 Boundary Class	<i>Boundary Class</i> , berisi kumpulan kelas yang menjadi <i>interfaces</i> atau interaksi antara satu atau lebih aktor dengan sistem, seperti tampilan <i>form entry</i> dan form cetak

 Control Class	<i>Control Class</i> , suatu objek yang berisi logika aplikasi yang tidak memiliki tanggung jawab kepada entitas, contohnya adalah kalkulasi dan aturan bisnis yang melibatkan berbagai objek
 1: Message0()	<i>Message</i> , simbol mengirim pesan antar kelas
<i>Recursive</i> 	<i>Recursive</i> , menggambarkan pengiriman pesan yang dikirim untuk dirinya sendiri
<i>Activation</i> 	<i>Activation</i> , mewakili sebuah eksekusi operasi dari objek panjang kotak ini berbanding lurus dengan durasi aktivitas sebuah operasi
<i>Lifeline</i> 	<i>Lifeline</i> , garis titik-titik yang terhubung dengan objek sepanjang <i>lifeline</i> terdapat <i>activation</i>

2.8 Penelitian Terdahulu

Penelitian ini disusun berdasarkan sejumlah studi terdahulu yang dijadikan sebagai landasan referensi, guna mengidentifikasi persamaan serta menggali ide-ide baru. Berikut ini disajikan beberapa tinjauan dari hasil-hasil penelitian yang memiliki keterkaitan dengan topik yang sedang dikaji dalam studi ini.

1. Menurut penelitian (Oktha Pratiwi et al., 2022) dengan judul **“Klasifikasi Penentuan Penerima Bantuan Sosial Tunai (Bst) Menggunakan Algoritma C4.5 Di Desa Keramas, Gianyar Bali”**. Dalam penelitian ini menyatakan bahwa menggunakan algoritma C4.5 untuk mengklasifikasikan

penerima bantuan sosial tunai (BST) di Desa Keramas, Gianyar Bali. Algoritma C4.5 dipilih karena kemampuannya dalam menangani atribut data yang kompleks dan menghasilkan pohon keputusan yang dapat membantu proses pengambilan keputusan secara objektif dan transparan.

2. Penelitian yang dikemukakan oleh (Muhamad Adhi Mukti et al., 2023) pada judul **“Implementasi Algoritma Naive Bayes untuk Prediksi Calon Penerima Blt (Studi Kasus Desa Tridayasakti, Kabupaten Bekasi)”**. Dalam penelitian ini menyatakan bahwa Penelitian ini mengimplementasikan algoritma Naive Bayes untuk memprediksi calon penerima Bantuan Langsung Tunai (BLT) di Desa Tridayasakti, Kabupaten Bekasi. Dengan memanfaatkan data historis penerima BLT dan atribut-atribut yang relevan, algoritma Naive Bayes menunjukkan akurasi yang baik dalam melakukan klasifikasi, sehingga dapat membantu dalam proses seleksi penerima bantuan yang lebih efektif dan efisien.
3. Berdasarkan penelitian yang dilakukan oleh (Isyam et al., 2024) yang memiliki judul **“Klasifikasi Penerima Bantuan Program Keluarga Harapan Menggunakan Support Vector Machine”**. Dalam penelitian ini, algoritma Support Vector Machine (SVM) diterapkan untuk mengklasifikasikan penerima Bantuan Program Keluarga Harapan. Metode SVM dipilih karena kemampuannya dalam menangani data dengan dimensi tinggi dan menyediakan klasifikasi yang robust, sehingga dapat meningkatkan akurasi dalam menentukan kelayakan penerima bantuan berdasarkan karakteristik sosial dan ekonomi yang relevan.

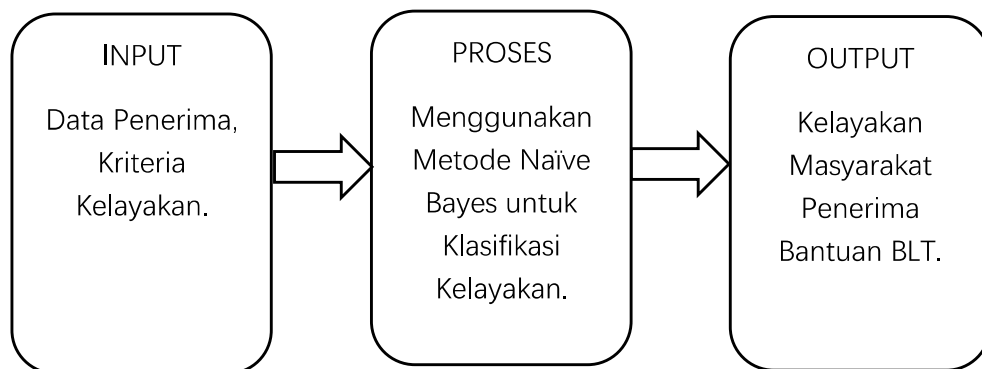
4. Penelitian yang dilakukan oleh (Aulia, 2021) berjudul **“Klasterisasi Pola Penjualan Pestisida Menggunakan Metode K-means Clustering (Studi Kasus Di Toko Juanda Tani Kecamatan Hutabayu Raja)”**. Penelitian ini menggunakan metode K-Means untuk menganalisis pola penjualan pestisida di Toko Juanda Tani, Kecamatan Hutabayu Raja. Dengan menerapkan algoritma K-Means, penelitian ini berhasil mengelompokkan data penjualan berdasarkan karakteristik produk dan perilaku konsumen, sehingga memberikan wawasan yang lebih baik dalam pengambilan keputusan terkait strategi pemasaran dan persediaan.
5. Pada penelitian yang dilakukan oleh (Teguh Adrian & Nana Suarna, 2024) dengan judul **“Implementasi Data Mining Untuk Mengklasifikasikan Hasil Kelulusan Madrasah Menggunakan Algoritma Naïve Bayes”**. Penelitian ini menggunakan metode K-Means Clustering untuk menganalisis pola penjualan pestisida di Toko Juanda Tani, Kecamatan Hutabayu Raja. Penelitian ini menggunakan algoritma K-Means untuk mengelompokkan data penjualan berdasarkan karakteristik produk dan perilaku konsumen. Ini menawarkan wawasan yang lebih baik tentang strategi pemasaran dan persediaan.
6. Pada penelitian yang dikemukakan oleh (Sholeh et al., 2023) dengan judul **“Sistem Pendukung Keputusan Penentuan Penerima BLT-DD Menggunakan Algoritma C4.5”**. Implementasi algoritma klasifikasi mampu memberikan pendekatan yang efisien untuk mendukung proses pengambilan keputusan. Tujuan utama dari metode naïve bayes sejalan dengan hal ini,

yakni memberikan rekomendasi secara cepat dan akurat dalam mengevaluasi kelayakan calon penerima bantuan.

7. Berdasarkan penelitian oleh (Kurniadi et al., 2022) berjudul **“KLASIFIKASI MASYARAKAT PENERIMA BANTUAN LANGSUNG TUNAI DADES A MENGGUNAKAN NAÏVE BAYES DAN SMOTE”**.

Untuk mengklasifikasikan data keluarga yang termasuk dalam kategori layak dan tidak layak menerima BLT Dana Desa (DD), penelitian ini menggunakan metode Naive Bayes. Disamping itu, guna menangani permasalahan ketidakseimbangan kelas pada data yang dianalisis, studi ini menerapkan metode *Synthetic Minority Over-sampling Technique* (SMOTE) sebagai solusi.

2.9 Kerangka Pemikiran



Kerangka pemikiran merupakan alur logis yang menghubungkan antara teori atau konsep dengan variabel-variabel penelitian yang telah diidentifikasi sebagai permasalahan, dan disusun dalam bentuk rancangan alur untuk penyelesaian penelitian. Rangkaian sistem yang diusulkan dijelaskan melalui tiga tahap utama, yaitu: Input, Proses, dan Output. Setiap tahapan ini memiliki peran

yang krusial dalam menilai kelayakan individu dalam mendapatkan dukungan sosial sesuai dengan kriteria yang telah ditetapkan.

a. Input

Pada tahap Input, data yang diperlukan untuk analisis dikumpulkan. Data Penerima mencakup informasi pribadi dari calon penerima bantuan, seperti nama, usia, status pekerjaan, pendapatan, dan jumlah tanggungan keluarga.

b. Proses

Setelah data berhasil dikumpulkan, langkah berikutnya adalah proses yang berfokus pada penerapan algoritma Naïve Bayes untuk mengklasifikasikan kelayakan penerima bantuan. Naïve Bayes adalah salah satu metode statistik dalam bidang pembelajaran mesin yang digunakan untuk melakukan klasifikasi dengan pendekatan probabilistik.

c. Output

Tahap akhir dari alur ini adalah Output, yang menghasilkan keputusan tentang Kelayakan Masyarakat Penerima Bantuan BLT. Output ini berisi hasil klasifikasi yang menunjukkan apakah calon penerima bantuan layak atau tidak untuk menerima bantuan berdasarkan data yang telah dianalisis.

objektif, dan efisien dalam menentukan kelayakan masyarakat sebagai penerima bantuan BLT. Pendekatan ini mendukung proses pengambilan keputusan yang lebih tepat dan sesuai dengan kriteria yang telah ditentukan oleh pemerintah.