

BAB III

METODOLOGI PENELITIAN

3.1. Metode Penelitian

Penelitian ini menggunakan metode kuantitatif dengan pendekatan deskriptif dan prediktif. Metode kuantitatif dipilih karena penelitian ini menganalisis data numerik yang diperoleh dari publikasi resmi Badan Pusat Statistik (BPS) Kabupaten Labuhanbatu, sehingga memungkinkan pengukuran dan analisis secara objektif serta terukur.

Pendekatan deskriptif diterapkan untuk menggambarkan kondisi serta tren konsumsi pangan masyarakat Kabupaten Labuhanbatu berdasarkan data pengeluaran konsumsi pangan per kapita menurut kelompok komoditas. Melalui pendekatan ini, penelitian menyajikan deskripsi perubahan pola konsumsi pangan dari waktu ke waktu secara sistematis dan terstruktur.

Sementara itu, pendekatan prediktif digunakan untuk memperkirakan kebutuhan pangan di masa depan dengan memanfaatkan metode *machine learning*. Pendekatan ini bertujuan memodelkan data historis konsumsi pangan sehingga menghasilkan estimasi kebutuhan pangan yang bersifat proyektif dan dapat dijadikan sebagai bahan pertimbangan dalam perencanaan kebijakan pangan di tingkat daerah.

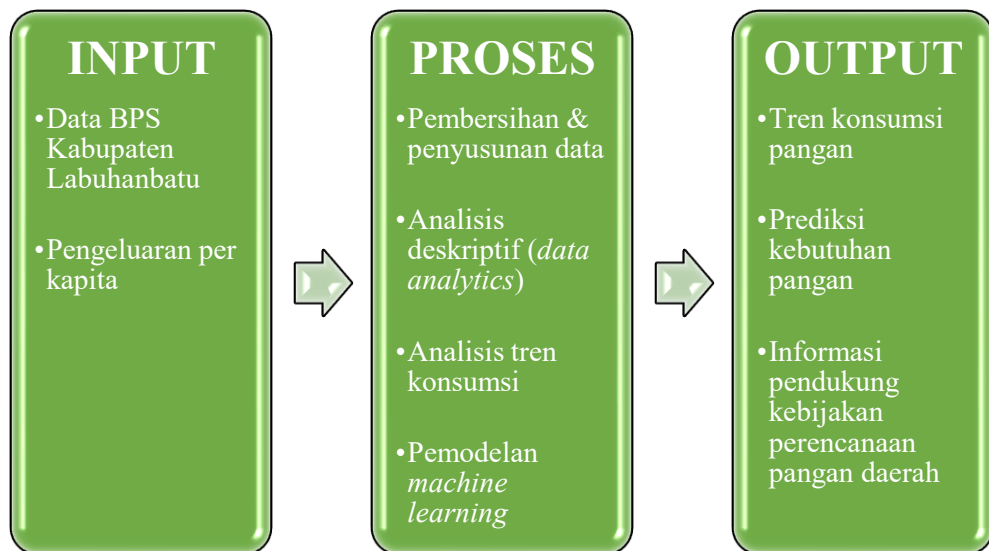
3.2. Diagram Alur Penelitian

Untuk mempermudah pemahaman mengenai tahapan dan alur penelitian, bagian ini menyajikan diagram alur penelitian yang menggambarkan proses penelitian secara sistematis. Diagram ini bertujuan memperjelas hubungan antara

input, proses, dan *output* penelitian, sekaligus menggambarkan keseluruhan tahapan pelaksanaan penelitian sebelum dijelaskan secara rinci mengenai lokasi dan waktu penelitian.

3.2.1. Diagram Blok Sistem (*Input – Proses – Output*)

Untuk mempermudah pemahaman terhadap alur penelitian, digunakan diagram blok sistem yang menggambarkan tahapan input, proses, dan output sebagaimana ditunjukkan pada Gambar 3.1.

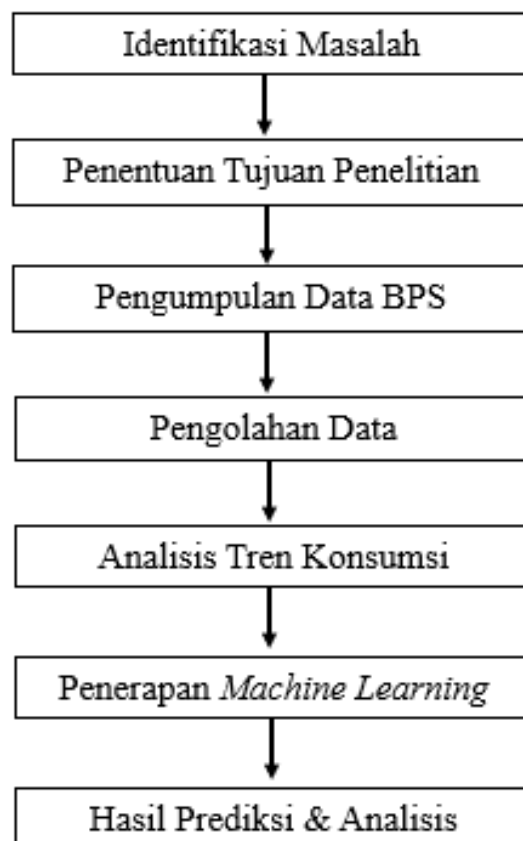


Gambar 3.1 Diagram Blok Sistem Penelitian (*Input–Proses–Output*)

Diagram blok sistem pada Gambar 3.1 menggambarkan bahwa data sekunder yang diperoleh dari publikasi resmi BPS berfungsi sebagai input utama penelitian. Data ini kemudian diproses melalui beberapa tahapan, meliputi analisis deskriptif, identifikasi tren konsumsi, dan pemodelan menggunakan metode *machine learning*. Hasil dari proses tersebut berupa informasi mengenai tren konsumsi pangan serta prediksi kebutuhan pangan, yang selanjutnya menjadi *output* utama penelitian.

3.2.2. Diagram Alir Penelitian (*Flowchart*)

Diagram alir penelitian pada Gambar 3.2 memperlihatkan tahapan penelitian secara sistematis dan berurutan, dimulai dari identifikasi masalah hingga tahap penyusunan laporan penelitian. Diagram ini memudahkan pemahaman mengenai alur kerja penelitian dan hubungan antar tahap, sehingga setiap proses dapat dilaksanakan secara terstruktur dan logis.

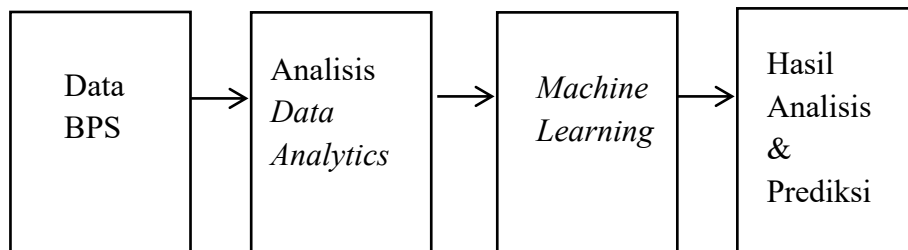


Gambar 3.2 Diagram Alir Penelitian (*Flowchart*)

3.2.3. Diagram Balok Penelitian

Diagram balok penelitian digunakan untuk menggambarkan tahapan utama dalam penelitian yang meliputi proses pengolahan data hingga

menghasilkan hasil analisis dan prediksi. Diagram balok penelitian dapat dilihat pada Gambar 3.3.



Gambar 3.3 Diagram Balok Penelitian

Diagram balok penelitian pada Gambar 3.3 menggambarkan hubungan antar komponen utama penelitian secara ringkas dan terstruktur. Data Badan Pusat Statistik dianalisis melalui pendekatan *data analytics* untuk mengidentifikasi pola dan tren konsumsi pangan, kemudian diproses lebih lanjut menggunakan metode *machine learning* guna menghasilkan prediksi kebutuhan pangan. Diagram ini mempermudah pemahaman alur penelitian secara visual, menunjukkan integrasi seluruh tahapan penelitian mulai dari data Badan Pusat Statistik sebagai sumber utama hingga rekomendasi kebijakan sebagai *output* akhir.

3.3. Tahapan Penelitian

Penjelasan tiap tahapan diberikan secara rinci untuk memperjelas proses penelitian dari awal hingga akhir.

1. Identifikasi Masalah

Tahap awal penelitian dimulai dengan mengidentifikasi permasalahan yang berkaitan dengan kebutuhan pangan dan tren konsumsi pangan masyarakat

Kabupaten Labuhanbatu, dengan mempertimbangkan kondisi aktual serta ketersediaan data statistik.

2. Penentuan Tujuan Penelitian

Setelah masalah teridentifikasi, tahap selanjutnya adalah merumuskan tujuan penelitian, yaitu menganalisis tren konsumsi pangan serta memprediksi kebutuhan pangan dengan menggunakan pendekatan *data analytics* dan metode *machine learning*.

3. Pengumpulan Data BPS

Data penelitian diperoleh dari publikasi resmi Badan Pusat Statistik Kabupaten Labuhanbatu, khususnya data pengeluaran konsumsi pangan per kapita menurut kelompok komoditas.

4. Pengolahan Data

Data yang telah dikumpulkan kemudian diseleksi, disusun, dan disiapkan untuk keperluan analisis. Tahap ini bertujuan memastikan konsistensi, keterbacaan, dan kualitas data yang digunakan.

5. Analisis Tren Konsumsi

Analisis tren konsumsi pangan dilakukan menggunakan pendekatan *data analytics* untuk mengidentifikasi pola perubahan konsumsi pangan masyarakat dari waktu ke waktu.

6. Penerapan *Machine Learning*

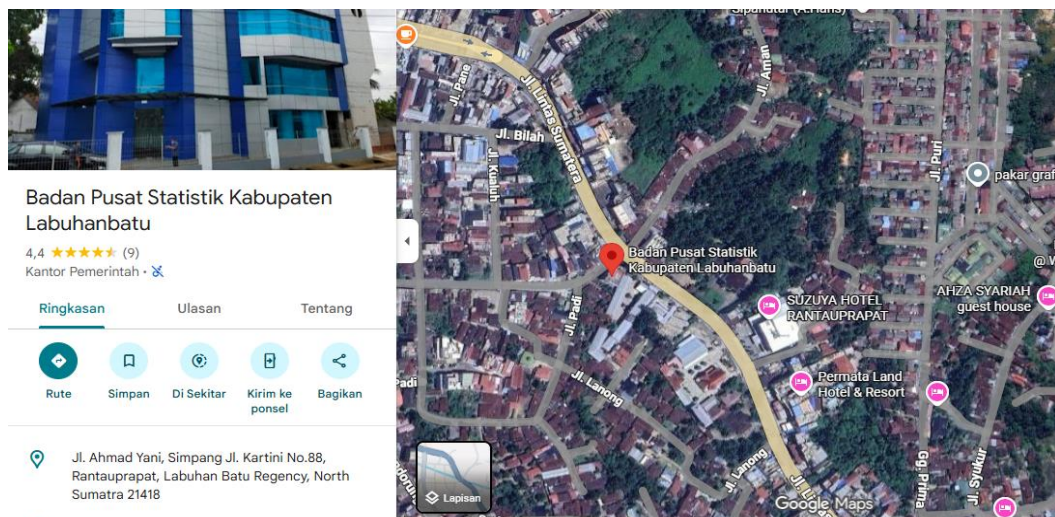
Pada tahap ini, data historis konsumsi pangan dimodelkan menggunakan metode *machine learning* guna menghasilkan prediksi kebutuhan pangan pada periode mendatang.

7. Hasil Prediksi & Analisis

Tahap akhir penelitian meliputi interpretasi hasil analisis dan prediksi, serta penyusunan laporan penelitian secara sistematis dan ilmiah.

3.4. Lokasi dan Waktu Penelitian

Penelitian ini dilaksanakan di Kabupaten Labuhanbatu, Provinsi Sumatera Utara. Pemilihan lokasi penelitian didasarkan pada ketersediaan data statistik resmi yang dipublikasikan oleh Badan Pusat Statistik (BPS) Kabupaten Labuhanbatu serta relevansinya dengan kajian kebutuhan dan konsumsi pangan daerah.



Gambar 3.4. Lokasi Penelitian

Waktu pelaksanaan penelitian direncanakan berlangsung selama kurang lebih 4 (empat) bulan yakni dari bulan Januari – April tahun 2026, mulai dari pengumpulan data, pengolahan data, analisis, hingga penyusunan laporan skripsi.

Untuk memperjelas pelaksanaan penelitian, berikut disajikan tabel rencana waktu penelitian.

Tabel 3.1. Waktu Penelitian

No.	Kegiatan	Tahun 2026			
		Januari	Februari	Maret	April
1.	Identifikasi masalah dan studi literatur				
2.	Pengumpulan data dari BPS				
3.	Pra-pemrosesan dan pembersihan data				
4.	Analisis tren konsumsi pangan				
5.	Pembangunan model <i>machine learning</i>				
6.	Evaluasi dan validasi model				
7.	Interpretasi hasil dan pembahasan				
8.	Penyusunan dan penyempurnaan laporan skripsi				

3.5. Jumlah dan Struktur Data Penelitian

Subbab ini menjelaskan jumlah dan struktur data yang digunakan dalam penelitian, dengan tujuan memberikan gambaran yang jelas mengenai cakupan data serta memastikan kecukupan informasi untuk analisis tren dan prediksi kebutuhan pangan.

Data penelitian merupakan data tahunan pengeluaran konsumsi pangan per kapita per bulan menurut kelompok komoditas, yang diperoleh dari publikasi resmi Badan Pusat Statistik Kabupaten Labuhanbatu. Periode data yang digunakan mencakup tahun 2020 hingga 2025. Dengan demikian, jumlah observasi ditentukan berdasarkan kombinasi antara jumlah tahun pengamatan dan jumlah kelompok komoditas pangan yang dianalisis, sehingga struktur data mencerminkan keseluruhan sampel yang representatif untuk penelitian ini.

Jumlah observasi yang digunakan dalam penelitian ini relatif terbatas karena data yang dianalisis bersifat tahunan dan agregat di tingkat kabupaten. Oleh karena itu, penelitian ini tidak menerapkan model *machine learning* kompleks seperti *deep learning*, melainkan menggunakan algoritma yang lebih sederhana, stabil, dan mudah diinterpretasikan. Pendekatan ini dipilih agar hasil prediksi tetap relevan, valid, dan sesuai dengan tujuan penelitian sebagai analisis tren dan prediksi awal kebutuhan pangan daerah.

Jumlah observasi yang digunakan dalam penelitian ini dianggap memadai untuk melakukan analisis tren konsumsi pangan serta pemodelan prediktif kebutuhan pangan dengan memanfaatkan pendekatan *data analytics* dan *machine learning* pada tingkat kabupaten. Pemilihan lokasi penelitian, yaitu Kabupaten Labuhanbatu, didasarkan pada ketersediaan data statistik resmi dari BPS serta

relevansinya dengan studi mengenai kebutuhan dan pola konsumsi pangan di tingkat daerah.

Periode penelitian mencakup tahapan pengumpulan data, pengolahan, analisis, hingga penyusunan laporan penelitian. Data yang digunakan bersifat tahunan, dengan rentang waktu 2020–2025, sesuai dengan publikasi resmi yang disediakan oleh BPS Kabupaten Labuhanbatu.

3.6. Jenis dan Sumber Data

Jenis data yang digunakan dalam penelitian ini adalah data sekunder. Data sekunder diperoleh dari publikasi resmi Badan Pusat Statistik, khususnya BPS Kabupaten Labuhanbatu. Data yang digunakan meliputi data pengeluaran konsumsi pangan per kapita per bulan menurut kelompok komoditas pangan.

Selain data konsumsi pangan, penelitian ini juga menggunakan data pendukung seperti jumlah penduduk Kabupaten Labuhanbatu, jumlah konsumsi kalori dan protein sehari per kapita untuk memberikan konteks dalam analisis kebutuhan pangan. Seluruh data yang digunakan bersifat agregat dan dipublikasikan secara resmi oleh BPS, sehingga memiliki tingkat reliabilitas dan validitas yang tinggi.

3.7. Teknik Pengumpulan Data

Teknik pengumpulan data yang digunakan dalam penelitian ini adalah metode dokumentasi, yaitu dengan mengumpulkan data sekunder yang berasal dari laporan dan publikasi resmi Badan Pusat Statistik Kabupaten Labuhanbatu.

Data yang diperoleh selanjutnya diseleksi dan disesuaikan dengan kebutuhan penelitian.

3.8. Variabel Penelitian

Dalam penelitian ini, kebutuhan pangan didefinisikan secara operasional sebagai nilai estimasi atau hasil prediksi pengeluaran konsumsi pangan per kapita per bulan berdasarkan kelompok komoditas pangan. Pengeluaran konsumsi pangan digunakan sebagai proksi kebutuhan pangan karena mencerminkan tingkat konsumsi dan kemampuan akses masyarakat terhadap pangan, sebagaimana disajikan dalam data resmi Badan Pusat Statistik. Dengan demikian, peningkatan nilai pengeluaran konsumsi pangan diinterpretasikan sebagai peningkatan kebutuhan pangan masyarakat. Dalam penelitian ini, kebutuhan pangan didefinisikan secara operasional sebagai nilai estimasi atau hasil prediksi pengeluaran konsumsi pangan per kapita per bulan berdasarkan kelompok komoditas pangan. Pengeluaran konsumsi pangan digunakan sebagai proksi kebutuhan pangan karena mencerminkan tingkat konsumsi dan kemampuan akses masyarakat terhadap pangan, sebagaimana disajikan dalam data resmi Badan Pusat Statistik. Dengan demikian, peningkatan nilai pengeluaran konsumsi pangan diinterpretasikan sebagai peningkatan kebutuhan pangan masyarakat.

Variabel penelitian merupakan objek yang diamati dan dianalisis dalam penelitian ini. Berdasarkan tujuan penelitian serta ketersediaan data dari BPS, variabel yang digunakan dibagi menjadi variabel dependen dan variabel independen sebagai berikut:

1. Variabel Independen (X)

Variabel independen meliputi:

- a. Tahun pengamatan (*time – series*), yang merepresentasikan periode data historis dari 2020 hingga 2025.
- b. Kelompok komoditas pangan, yang terdiri atas padi-padian, umbi-umbian, ikan/udang/kerang, daging serta telur dan susu.

2. Variabel Dependen (Y)

Variabel dependen dalam penelitian ini adalah kebutuhan pangan masyarakat Kabupaten Labuhanbatu, yang diprosikan melalui pengeluaran konsumsi pangan per kapita per bulan berdasarkan kelompok komoditas pangan.

Tabel 3.2 Variabel Penelitian

No	Variabel	Jenis	Keterangan
1	Tahun	Independen	Periode 2020 – 2025
2	Komoditas Pangan	Independen	5 kelompok pangan
3	Jumlah Penduduk	Independen	ribu jiwa
4	Konsumsi kalori	Independen	Kkal/hari
5	Konsumsi Protein	Independen	gr/hari
6	Pengeluaran Pangan	Dependen	Rp/kapita/bulan
7	Tren Konsumsi	Dependen	Naik, Turun, Stabil

3.9. Model dan Rumus Analisis yang Digunakan

Subbab ini menjelaskan model dan rumus analisis yang diterapkan dalam penelitian, disesuaikan dengan tujuan penelitian serta karakteristik data BPS yang bersifat agregat dan memiliki jumlah observasi terbatas. Oleh karena itu, analisis difokuskan pada analisis deskriptif, analisis tren, dan pemodelan prediktif

sederhana, yang mudah diinterpretasikan dan relevan untuk mendukung perencanaan pangan di tingkat daerah.

3.9.1. Rata – Rata (*Mean*)

Rata-rata digunakan dalam analisis deskriptif untuk menggambarkan nilai tengah pengeluaran konsumsi pangan per kapita menurut kelompok komoditas selama periode pengamatan. Rumus rata-rata yang digunakan adalah:

$$\bar{X} = \frac{\sum X_i}{n}$$

Keterangan:

\bar{X} = nilai rata-rata pengeluaran konsumsi pangan

X_i = nilai pengeluaran konsumsi pangan pada periode ke- i

n = jumlah periode pengamatan

3.9.2. Model Tren Linear

Analisis tren linear diterapkan untuk mengidentifikasi kecenderungan perubahan pengeluaran konsumsi pangan dari waktu ke waktu. Model tren linear dirumuskan sebagai berikut:

$$Y = a + bX$$

Keterangan:

Y = pengeluaran konsumsi pangan per kapita

X = waktu (tahun)

a = konstanta

b = koefisien tren

Nilai koefisien b menunjukkan arah perubahan konsumsi pangan, yaitu meningkat, menurun, atau relatif stabil selama periode pengamatan.

3.9.3. Model *Linear Regression*

Model *Linear Regression* digunakan untuk memperkirakan nilai kebutuhan pangan pada tahun selanjutnya dengan memanfaatkan variabel jumlah penduduk, konsumsi kalori, dan konsumsi protein. Metode ini diterapkan untuk menganalisis hubungan antara variabel independen dan variabel dependen sehingga dapat menghasilkan nilai prediksi berdasarkan hubungan antarvariabel tersebut.

Model regresi linear berganda dinyatakan sebagai berikut:

$$y = a + b_1x_1 + b_2x_2 + b_3x_3$$

Keterangan:

y = Total kebutuhan pangan (Variabel Dependen)

a = Konstanta (*intercept*)

b_1, b_2, b_3 = Koefisien regresi

x_1 = Jumlah Penduduk

x_2 = Konsumsi Kalori

x_3 = Konsumsi Protein

Model ini digunakan untuk menghitung nilai perkiraan total kebutuhan pangan berdasarkan nilai variabel independen yang tersedia pada setiap periode tahun. Dengan menggunakan model tersebut, hubungan antara variabel yang

memengaruhi dan variabel yang diprediksi dapat dianalisis untuk menghasilkan nilai estimasi.

Dalam penelitian ini, variabel independen yang digunakan terdiri dari jumlah penduduk, konsumsi kalori, dan konsumsi protein. Sementara itu, variabel dependen yang menjadi objek prediksi adalah total kebutuhan pangan..

3.9.4. Metode *Naïve Bayes*

Metode Naive Bayes diterapkan untuk mengklasifikasikan tren konsumsi pangan berdasarkan perubahan nilai total pengeluaran. Metode ini bekerja dengan memanfaatkan pendekatan probabilitas guna menentukan kemungkinan suatu data masuk ke dalam kategori atau kelas tertentu.

Naïve Bayes bekerja dengan menghitung peluang suatu data termasuk ke dalam suatu kelas berdasarkan nilai atribut yang dimiliki. Metode ini memiliki asumsi bahwa setiap variabel bersifat independen satu sama lain, sehingga perhitungan probabilitas menjadi lebih sederhana.

Model *Naïve Bayes* dinyatakan sebagai berikut:

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

Keterangan:

$P(C|X)$ = Probabilitas kelas berdasarkan data

$P(X|C)$ = Probabilitas data pada kelas tertentu

$P(C)$ = Probabilitas awal kelas

$P(X)$ = Probabilitas total data

Dalam penelitian ini, metode *Naïve Bayes* digunakan untuk mengklasifikasikan tren konsumsi pangan ke dalam tiga kategori, yaitu “Stabil”, “Naik” dan “Turun”. Proses klasifikasi dilakukan berdasarkan data pengeluaran pangan dari beberapa kelompok komoditas, seperti padi-padian, umbi-umbian, ikan, daging, serta susu dan telur.

Secara umum, prinsip kerja *Naïve Bayes* didasarkan pada Teorema Bayes yang menyatakan bahwa probabilitas suatu hipotesis dapat dihitung berdasarkan probabilitas awal dan bukti yang ada. Dalam konteks penelitian ini, hipotesis yang dimaksud adalah kategori tren konsumsi pangan.

Tahapan penerapan metode *Naïve Bayes* dalam penelitian ini meliputi:

1. Menentukan variabel *input* (atribut) dan variabel target (kelas).
2. Menghitung probabilitas setiap kelas berdasarkan data latih.
3. Menghitung probabilitas atribut terhadap masing-masing kelas.
4. Mengalikan seluruh probabilitas untuk menentukan kelas dengan nilai probabilitas tertinggi.
5. Menentukan hasil klasifikasi berdasarkan nilai probabilitas terbesar.

Implementasi metode *Naïve Bayes* dalam penelitian ini dilakukan menggunakan perangkat lunak *RapidMiner*. Data yang telah diproses sebelumnya diimpor ke dalam *RapidMiner*, kemudian dilakukan pemisahan data menjadi data latih (*training*) dan data uji (*testing*). Selanjutnya, model *Naïve Bayes* dibangun menggunakan data latih dan diuji menggunakan data uji untuk mengetahui tingkat akurasi model.

Pemilihan metode *Naïve Bayes* dalam penelitian ini didasarkan pada keunggulannya yang sederhana, cepat, dan mampu bekerja dengan baik pada

dataset berukuran kecil. Namun demikian, metode ini memiliki keterbatasan, yaitu asumsi independensi antar variabel yang tidak selalu sesuai dengan kondisi data nyata.

Dengan demikian, metode *Naïve Bayes* diharapkan dapat memberikan gambaran mengenai kemampuan model dalam mengklasifikasikan tren konsumsi pangan berdasarkan data yang tersedia.

3.10. Implementasi *Machine Learning* Menggunakan *RapidMiner*

Implementasi *machine learning* dalam penelitian ini dilakukan dengan memanfaatkan perangkat lunak *RapidMiner Studio*. Perangkat lunak tersebut digunakan sebagai alat untuk membantu proses pengolahan data, pembangunan model prediksi, serta analisis tren konsumsi pangan. Tahap implementasi dilakukan dengan menggunakan data historis konsumsi pangan dari tahun 2020 hingga 2025 yang telah disusun dalam bentuk tabel menggunakan *Microsoft Excel*.

Data yang digunakan dalam penelitian ini mencakup beberapa variabel, yaitu jumlah penduduk, konsumsi kalori, konsumsi protein, serta total kebutuhan pangan. Sebelum dilakukan analisis, data terlebih dahulu disiapkan dalam format *Microsoft Excel*. Selanjutnya, data tersebut diimpor ke dalam *RapidMiner* untuk diproses lebih lanjut menggunakan metode *machine learning* yang digunakan dalam penelitian ini.

3.10.1 *Import Data ke RapidMiner*

Tahap awal dalam implementasi *machine learning* adalah melakukan proses impor data dari *Microsoft Excel* ke dalam *RapidMiner Studio*. Data yang telah disusun dalam bentuk tabel terlebih dahulu disimpan dalam format *.xlsx*, kemudian dimasukkan ke dalam *RapidMiner* dengan menggunakan operator *Retrieve*.

Langkah-langkah yang dilakukan pada tahap ini meliputi:

1. Membuka aplikasi *RapidMiner*.
2. Memilih menu *Import Data* untuk memasukkan data yang berasal dari file *Excel*.
3. Menentukan file data yang telah dipersiapkan sebelumnya.
4. Memastikan setiap atribut pada dataset telah dikenali dengan benar sesuai dengan tipe datanya.
5. Menyimpan *dataset* tersebut ke dalam *repository RapidMiner*.

Setelah proses impor data selesai dilakukan, *dataset* yang telah tersimpan di *repository RapidMiner* dapat digunakan untuk tahap selanjutnya, yaitu pembangunan model *machine learning*.

3.10.2 *Implementasi Model Linear Regression*

Metode *Linear Regression* digunakan dalam penelitian ini untuk memperkirakan total kebutuhan pangan pada tahun berikutnya. Model prediksi dibangun dengan menggunakan beberapa variabel independen, yaitu jumlah

penduduk, konsumsi kalori, dan konsumsi protein. Sementara itu, variabel dependen yang diprediksi dalam penelitian ini adalah total kebutuhan pangan.

Proses implementasi model *Linear Regression* pada perangkat lunak *RapidMiner Studio* dilakukan melalui beberapa tahapan sebagai berikut:

1. Menambahkan *operator Retrieve* untuk mengambil *dataset* yang telah tersimpan pada *repository*.
2. Menggunakan *operator Select Attributes* untuk menentukan atribut yang akan digunakan dalam proses analisis.
3. Menggunakan *operator Set Role* untuk menetapkan atribut Total Kebutuhan Pangan sebagai label atau variabel yang akan diprediksi.
4. Menambahkan *operator Linear Regression* untuk membangun model prediksi berdasarkan data yang tersedia.
5. Menggunakan *operator Apply Model* untuk menghasilkan nilai prediksi dari model yang telah dibangun.
6. Menjalankan proses (*Run*) untuk memperoleh hasil prediksi kebutuhan pangan.

Hasil dari proses tersebut berupa nilai prediksi total kebutuhan pangan yang selanjutnya digunakan untuk memperkirakan kebutuhan pangan pada periode berikutnya, yaitu tahun 2026.

3.10.3 Implementasi Model *Naïve Bayes* untuk Analisis Tren

Metode *Naïve Bayes* digunakan dalam penelitian ini untuk menganalisis tren konsumsi pangan berdasarkan perubahan nilai total pengeluaran dari tahun ke

tahun. Sebelum proses klasifikasi dilakukan, terlebih dahulu dibentuk variabel tren yang terdiri dari tiga kategori, yaitu naik, turun, dan stabil.

Penentuan kategori tren dilakukan dengan membandingkan nilai total pengeluaran pada setiap tahun. Apabila nilai total pada tahun berjalan lebih tinggi dibandingkan tahun sebelumnya, maka tren dikategorikan sebagai naik. Sebaliknya, jika nilainya lebih rendah maka dikategorikan sebagai turun. Sementara itu, apabila nilai total pengeluaran pada dua tahun berturut-turut memiliki nilai yang sama, maka tren dikategorikan sebagai stabil.

Proses implementasi metode *Naive Bayes* pada perangkat lunak *RapidMiner Studio* dilakukan melalui beberapa tahapan berikut:

1. Menambahkan *operator Retrieve* untuk mengambil *dataset* yang berisi data tren.
2. Menggunakan *operator Set Role* untuk menetapkan atribut Tren sebagai label atau kelas yang akan diklasifikasikan.
3. Menambahkan *operator Naive Bayes* untuk membangun model klasifikasi berdasarkan data yang tersedia.
4. Menggunakan *operator Apply Model* untuk menghasilkan hasil klasifikasi tren dari model yang telah dibuat.
5. Menambahkan *operator Performance (Classification)* untuk mengevaluasi tingkat akurasi model klasifikasi.
6. Menjalankan proses (*Run*) untuk memperoleh hasil klasifikasi tren konsumsi pangan.

Hasil dari proses tersebut berupa kategori tren konsumsi pangan yang menggambarkan kecenderungan perubahan pola konsumsi masyarakat dari waktu ke waktu.

3.11. Evaluasi Model

Evaluasi model dilakukan untuk menilai tingkat kinerja model *machine learning* yang telah dibangun dalam penelitian ini. Proses evaluasi dilakukan terhadap model *Linear Regression* dan *Naïve Bayes* dengan menggunakan metode pengukuran yang disesuaikan dengan jenis model yang diterapkan. Tahap evaluasi tersebut dilakukan dengan memanfaatkan operator *Performance* pada perangkat lunak *RapidMiner Studio*.

3.11.1 Evaluasi Model *Linear Regression*

Evaluasi pada model *Linear Regression* dilakukan untuk mengetahui tingkat perbedaan atau kesalahan antara nilai aktual dan nilai hasil prediksi. Dalam penelitian ini, metode evaluasi yang digunakan adalah *Mean Absolute Error* dan *Mean Squared Error*.

Mean Absolute Error (MAE) digunakan untuk mengukur rata-rata nilai kesalahan absolut antara nilai prediksi dan nilai aktual yang dihasilkan oleh model. Adapun rumus *Mean Absolute Error* (MAE) yang digunakan dalam penelitian ini adalah sebagai berikut:

$$MAE = \frac{\sum |Y_i - \hat{Y}_i|}{n}$$

Keterangan:

MAE = *Mean Absolute Error*

Y_i = nilai aktual pada periode ke i

\hat{Y}_i = nilai hasil prediksi pada periode ke $- i$

n = jumlah data observasi

Mean Absolute Error (MAE) digunakan untuk menghitung rata-rata selisih absolut antara nilai aktual dan nilai hasil prediksi. Semakin kecil nilai MAE yang dihasilkan, maka semakin baik model prediksi yang dibangun. Selain MAE, digunakan juga metode *Mean Squared Error* (MSE) untuk mengukur tingkat kesalahan model.

Rumus *Mean Squared Error* (MSE) adalah sebagai berikut:

$$MSE = \frac{\sum (Y_i - \hat{Y}_i)^2}{n}$$

Keterangan:

MSE = *Mean Squared Error*

Y_i = nilai aktual pada periode ke $- i$

\hat{Y}_i = nilai hasil prediksi pada periode ke $- i$

n = jumlah data observasi

Mean Squared Error digunakan untuk menghitung rata-rata kuadrat dari selisih antara nilai aktual dengan nilai hasil prediksi yang dihasilkan oleh model. Semakin kecil nilai MSE yang diperoleh, maka menunjukkan bahwa tingkat kesalahan model semakin rendah.

Dalam penelitian ini, perhitungan nilai *Mean Absolute Error* dan *Mean Squared Error* (MSE) dilakukan secara otomatis dengan menggunakan operator *Performance (Regression)* pada perangkat lunak *RapidMiner Studio*.

3.11.2 Evaluasi Model *Naïve Bayes*

Evaluasi pada model *Naïve Bayes* dilakukan untuk mengetahui tingkat ketepatan dalam proses klasifikasi tren konsumsi pangan. Metode evaluasi yang digunakan adalah *Accuracy*, yaitu ukuran yang menunjukkan persentase jumlah data yang berhasil diklasifikasikan dengan benar oleh model.

Nilai *accuracy* diperoleh dengan membandingkan jumlah prediksi yang benar terhadap keseluruhan data yang digunakan dalam proses pengujian model. Semakin tinggi nilai *accuracy* yang dihasilkan, maka semakin baik pula kemampuan model dalam mengklasifikasikan tren konsumsi pangan.

Selain menggunakan *accuracy*, evaluasi model *Naïve Bayes* juga dilakukan dengan menganalisis *Confusion Matrix*. *Confusion matrix* digunakan untuk menunjukkan jumlah data yang diklasifikasikan secara benar maupun salah pada masing-masing kategori tren, yaitu naik, turun, dan stabil.

3.12. *Tools* dan Perangkat Lunak

Pemilihan *tools* dan perangkat lunak dalam penelitian ini didasarkan pada kesesuaiannya dengan tujuan penelitian, karakteristik data yang digunakan, serta kemudahan dalam proses pengolahan dan analisis data. Penggunaan perangkat lunak yang tepat diharapkan dapat meningkatkan ketepatan hasil analisis, efisiensi waktu penelitian, serta menjamin keterulangan (*reproducibility*) hasil penelitian. Adapun *tools* dan perangkat lunak yang digunakan dalam penelitian ini meliputi *Microsoft Excel* dan *RapidMiner*.

3.12.1. *Microsoft Excel*

Microsoft Excel dimanfaatkan pada tahap awal pengolahan data dalam proses *data analytics*, khususnya untuk pengolahan data dasar dan analisis deskriptif. Data pengeluaran konsumsi pangan per kapita yang diperoleh dari publikasi resmi Badan Pusat Statistik Kabupaten Labuhanbatu terlebih dahulu dimasukkan dan disusun dalam *Microsoft Excel*. Pada tahap ini, *Excel* digunakan untuk melakukan pembersihan data (*data cleaning*), pemeriksaan kelengkapan data, penyesuaian format data deret waktu (*time series*), serta perhitungan statistik deskriptif seperti nilai rata-rata, nilai minimum, dan nilai maksimum. Selain itu, *Microsoft Excel* juga digunakan untuk menyajikan data dalam bentuk tabel dan grafik tren konsumsi pangan guna mempermudah pemahaman terhadap pola perubahan konsumsi dari waktu ke waktu.

Penggunaan *Microsoft Excel* dalam penelitian ini dinilai sesuai karena data yang dianalisis bersifat agregat dengan jumlah observasi yang relatif terbatas. *Excel* menyediakan fitur pengolahan data yang sederhana namun cukup andal untuk mendukung analisis deskriptif dan eksploratif, sehingga hasil analisis awal dapat menjadi landasan yang kuat sebelum dilakukan tahap pemodelan prediktif.

3.12.2. *RapidMiner*

RapidMiner digunakan sebagai perangkat lunak utama dalam penerapan metode *machine learning* untuk memprediksi kebutuhan pangan masyarakat Kabupaten Labuhanbatu. *RapidMiner* merupakan salah satu *tools data science* yang menyediakan berbagai algoritma *machine learning* dan analisis data dalam lingkungan visual berbasis *workflow*. Keunggulan *RapidMiner* terletak

pada kemampuannya dalam membangun dan menguji model *machine learning* tanpa memerlukan pemrograman secara langsung, sehingga memudahkan proses eksperimen dan evaluasi model.

Dalam penelitian ini, *RapidMiner* dimanfaatkan untuk membangun model regresi deret waktu berbasis *machine learning* dengan menggunakan data historis pengeluaran konsumsi pangan. Tahapan pemodelan meliputi proses *input* data, pemilihan algoritma, pelatihan model (*training*), serta evaluasi hasil prediksi menggunakan metrik evaluasi seperti *Mean Absolute Error* (MAE) dan *Mean Squared Error* (MSE). Selain itu, *RapidMiner* juga menyediakan fasilitas visualisasi hasil prediksi yang membantu dalam interpretasi serta analisis kinerja model.

Pemilihan *RapidMiner* sebagai *tools machine learning* dinilai tepat karena karakteristik data penelitian ini berupa data *time series* dengan jumlah observasi yang terbatas. *RapidMiner* memungkinkan penerapan model yang relatif sederhana namun tetap memiliki tingkat akurasi yang baik dan mudah untuk diinterpretasikan, sehingga hasil prediksi yang dihasilkan dapat dipahami oleh peneliti maupun pihak nonteknis, seperti pengambil kebijakan.

Dengan demikian, kombinasi penggunaan *Microsoft Excel* dan *RapidMiner* dalam penelitian ini membentuk alur kerja yang terstruktur, dimulai dari tahap pengolahan dan analisis data awal hingga pemodelan serta evaluasi prediksi kebutuhan pangan. Pendekatan ini mendukung pencapaian tujuan penelitian dalam menghasilkan analisis tren konsumsi dan prediksi kebutuhan pangan yang akurat, efisien, dan berbasis data.

Dengan metodologi penelitian yang terstruktur dan berbasis data statistik resmi, penelitian ini diharapkan mampu menghasilkan analisis tren konsumsi dan prediksi kebutuhan pangan yang valid serta dapat dipertanggungjawabkan secara ilmiah. Pendekatan ini diharapkan memberikan kontribusi nyata bagi pengembangan kajian pangan berbasis data di tingkat daerah.

3.13. Output Penelitian

Output penelitian berupa tabel dan grafik hasil nilai prediksi pengeluaran kebutuhan pangan tahun berikutnya, hasil klasifikasi tren konsumsi, serta hasil evaluasi dari pengeluaran kebutuhan pangan masyarakat Kabupaten Labuhanbatu.