

## BAB II

### TINJAUAN PUSTAKA

#### 2.1. Penelitian Terdahulu yang Relevan

Penelitian terdahulu memiliki peran penting dalam menentukan posisi penelitian yang dilakukan serta mengidentifikasi celah penelitian (*research gap*) yang masih dapat dikembangkan. Selain itu, telaah terhadap studi sebelumnya juga berfungsi untuk memperkuat landasan teoritis dan metodologis yang digunakan dalam penelitian ini.

Zhou et al. (2022) melakukan penelitian terkait prediksi konsumsi pangan dengan memanfaatkan pendekatan *machine learning* pada data deret waktu. Hasil penelitian tersebut menunjukkan bahwa model *machine learning* mampu mengidentifikasi pola nonlinier dalam data konsumsi pangan serta menghasilkan tingkat akurasi yang lebih baik dibandingkan metode statistik konvensional. Temuan ini menunjukkan bahwa *machine learning* memiliki potensi yang signifikan dalam menganalisis pola konsumsi pangan yang bersifat dinamis.

Singh dan Kumar (2023) mengkaji prediksi permintaan pangan berbasis data pengeluaran rumah tangga menggunakan algoritma *Random Forest*. Penelitian ini membuktikan bahwa *Random Forest* mampu menurunkan tingkat kesalahan prediksi, khususnya pada data yang memiliki tingkat fluktuasi yang tinggi. Namun demikian, penelitian tersebut belum memanfaatkan data resmi Badan Pusat Statistik (BPS) daerah sebagai sumber data utama.

Dalimunthe et al. (2023) melakukan studi literatur mengenai penerapan *machine learning* dalam *demand forecasting*. Hasil kajian menunjukkan bahwa metode *machine learning* semakin banyak diterapkan karena kemampuannya

dalam memodelkan data historis yang kompleks dan beragam. Meskipun demikian, sebagian besar penelitian yang dikaji masih bersifat konseptual dan belum banyak diaplikasikan pada data lokal di tingkat kabupaten.

Phumchusri dan Sirimak (2024) mengembangkan model hibrida yang mengkombinasikan pendekatan analisis deret waktu dan *machine learning* dalam memprediksi permintaan pangan di Thailand. Penelitian tersebut menunjukkan bahwa penggabungan kedua pendekatan mampu meningkatkan akurasi prediksi secara signifikan dibandingkan dengan penggunaan satu metode secara terpisah.

Berdasarkan hasil telaah terhadap penelitian-penelitian terdahulu tersebut, dapat disimpulkan bahwa kajian yang secara khusus membahas prediksi kebutuhan pangan dengan memanfaatkan data resmi BPS pada tingkat kabupaten masih relatif terbatas. Oleh karena itu, penelitian ini memiliki unsur kebaruan dengan mengintegrasikan pendekatan *machine learning* dan *data analytics* pada data BPS Kabupaten Labuhanbatu.

Kebaruan (*novelty*) dari penelitian ini terletak pada penerapan metode *machine learning* berbasis regresi deret waktu (*time series regression*) untuk memprediksi kebutuhan pangan menggunakan data pengeluaran konsumsi pangan per kapita yang bersumber dari publikasi resmi Badan Pusat Statistik (BPS) pada tingkat kabupaten, khususnya Kabupaten Labuhanbatu. Berbeda dengan penelitian sebelumnya yang umumnya menggunakan data nasional atau provinsi, penelitian ini memfokuskan analisis pada data agregat daerah dengan jumlah observasi terbatas. Selain itu, penelitian ini mengintegrasikan pendekatan *data analytics* untuk analisis tren konsumsi pangan dan *machine learning* untuk prediksi kebutuhan pangan secara simultan, sehingga menghasilkan informasi

yang tidak hanya bersifat deskriptif tetapi juga prediktif dan relevan bagi perencanaan pangan daerah.

Tabel 2.1 Ringkasan Penelitian Terdahulu Terkait Prediksi Konsumsi Pangan

No.	Peneliti & Tahun	Data yang Digunakan	Metode	Hasil Utama	Perbedaan dengan Penelitian Ini
1	Zhou et al. (2022)	Data konsumsi pangan nasional	<i>Machine Learning</i>	Akurasi prediksi lebih tinggi dibanding regresi konvensional	Skala nasional
2	Singh & Kumar (2023)	Data pengeluaran rumah tangga	<i>Random Forest</i>	Error rendah	Tidak berbasis data BPS daerah
3	Dalimunthe et al. (2023)	Literatur	<i>Review ML</i>	ML efektif untuk demand forecasting	Studi konseptual
4	Phumchusri & Sirimak (2024)	Permintaan pangan Thailand	<i>Hybrid TS – ML</i>	Akurasi meningkat	Konteks luar negeri
5	Penelitian ini	Data BPS Kabupaten Labuhanbatu	<i>Machine Learning &amp; Data Analytics</i>	Prediksi tren dan kebutuhan pangan daerah	Fokus tingkat kabupaten dan data agregat

## **2.2. Kebutuhan Pangan dan Ketahanan Pangan**

### **2.2.1 Kebutuhan Pangan**

Pangan merupakan kebutuhan fundamental manusia yang harus dipenuhi secara berkelanjutan untuk menjamin kelangsungan hidup serta kualitas sumber daya manusia. Pemenuhan kebutuhan pangan berpengaruh langsung terhadap kesehatan, produktivitas, dan kualitas sumber daya manusia. Oleh karena itu, pangan menjadi salah satu indikator utama kesejahteraan masyarakat. Konsep kebutuhan pangan tidak hanya mencakup aspek kuantitatif, tetapi juga meliputi aksesibilitas, kestabilan ketersediaan dan pemanfaatan pangan yang efektif.

Dalam perspektif pembangunan daerah, pemenuhan kebutuhan pangan memiliki peran strategis karena berkaitan langsung dengan kesejahteraan masyarakat. Ketidakmampuan suatu wilayah dalam merencanakan dan mengelola kebutuhan pangan dapat menimbulkan kerentanan terhadap pangan, fluktuasi harga serta kesenjangan sosial-ekonomi. Oleh karena itu, analisis kebutuhan pangan perlu dilakukan secara sistematis dan berbasis data agar kebijakan yang dihasilkan dapat lebih tepat sasaran dan efektif.

Dalam penelitian ini, kebutuhan pangan didefinisikan sebagai kebutuhan ekonomi pangan, yang diproksikan melalui nilai pengeluaran konsumsi pangan per kapita. Pendekatan ini lazim digunakan dalam studi ekonomi konsumsi karena mencerminkan tingkat akses dan kemampuan rumah tangga dalam memenuhi kebutuhan pangannya. Meskipun tidak secara langsung mengukur konsumsi fisik atau kandungan gizi, pengeluaran konsumsi pangan dianggap representatif dalam menggambarkan kebutuhan pangan secara agregat, khususnya pada data statistik resmi seperti yang disediakan oleh Badan Pusat Statistik.

### **2.2.2 Ketahanan Pangan**

*Food and Agriculture Organization* (FAO) mendefinisikan ketahanan pangan sebagai kondisi terpenuhinya pangan bagi seluruh penduduk dan di mana seluruh penduduk memiliki akses terhadap pangan yang cukup, berkualitas, aman dan terjangkau secara konsisten sepanjang waktu.

Ketahanan pangan mencakup empat pilar utama, yaitu ketersediaan, akses, pemanfaatan, dan stabilitas pangan. Di tingkat daerah, ketahanan pangan sangat dipengaruhi oleh pola konsumsi masyarakat serta kemampuan pemerintah daerah dalam merencanakan kebutuhan pangan secara tepat berdasarkan data.

## **2.3. Pola dan Tren Konsumsi Pangan**

### **2.3.1 Pola Konsumsi Pangan**

Pola konsumsi pangan mencerminkan jenis, jumlah serta nilai konsumsi pangan yang dikonsumsi masyarakat dalam periode tertentu. Konsumsi pangan tidak hanya berkaitan dengan jumlah makanan yang dikonsumsi, tetapi juga kualitas gizi yang terkandung di dalamnya.

Konsumsi pangan masyarakat dipengaruhi oleh berbagai faktor, seperti tingkat pendapatan, harga pangan, serta ketersediaan bahan makanan. Data konsumsi pangan di Indonesia umumnya diperoleh dari survei sosial ekonomi yang dilakukan oleh Badan Pusat Statistik (BPS, 2022).

Selain itu, perubahan pola konsumsi pangan juga dipengaruhi oleh perkembangan ekonomi dan sosial masyarakat. Masyarakat dengan tingkat pendapatan yang lebih tinggi cenderung memiliki pola konsumsi yang lebih

beragam dibandingkan dengan masyarakat berpendapatan rendah (Pan et al., 2022).

Selain itu, dalam beberapa tahun terakhir, faktor eksternal seperti pandemi COVID-19 juga turut memengaruhi daya beli dan perilaku konsumsi masyarakat. Ketersediaan pangan di suatu daerah juga menjadi faktor penting dalam menentukan pola konsumsi masyarakat. Daerah yang memiliki produksi pangan yang tinggi cenderung memiliki tingkat konsumsi yang lebih stabil dibandingkan dengan daerah yang bergantung pada pasokan dari luar (Nigus & Dorsey, 2021).

Perubahan pola konsumsi pangan mencerminkan perubahan gaya hidup dan kondisi sosial ekonomi masyarakat. Oleh karena itu, analisis pola konsumsi pangan penting dilakukan untuk memahami kebutuhan pangan masyarakat. Dengan demikian, analisis kebutuhan pangan harus mempertimbangkan berbagai faktor yang saling berkaitan agar dapat memberikan gambaran yang lebih akurat.

### **2.3.2 Tren Konsumsi Pangan**

Tren konsumsi pangan menunjukkan arah perubahan pola konsumsi dari waktu ke waktu. Analisis tren konsumsi menjadi penting untuk mengidentifikasi kecenderungan perubahan kebutuhan pangan masyarakat, baik berupa peningkatan maupun penurunan konsumsi pada kelompok komoditas tertentu. Data dan informasi mengenai tren konsumsi ini menjadi landasan penting bagi perencanaan kebutuhan pangan jangka menengah maupun jangka panjang, terutama pada tingkat daerah.

#### **2.4. Pengaruh Jumlah Penduduk terhadap Konsumsi Pangan**

Jumlah penduduk merupakan salah satu faktor utama yang mempengaruhi kebutuhan pangan. Semakin besar jumlah penduduk, maka semakin tinggi pula kebutuhan konsumsi pangan yang harus dipenuhi.

Pertumbuhan penduduk yang tidak diimbangi dengan peningkatan produksi pangan dapat menyebabkan ketidakseimbangan antara kebutuhan dan ketersediaan pangan. Hal ini berpotensi menimbulkan permasalahan dalam ketahanan pangan suatu daerah (BPS, 2022).

Selain itu, struktur demografi juga mempengaruhi pola konsumsi masyarakat. Kelompok usia produktif memiliki kebutuhan energi yang lebih tinggi dibandingkan kelompok usia lainnya, sehingga mempengaruhi tingkat konsumsi pangan (Pan et al., 2022).

Dalam konteks daerah, peningkatan jumlah penduduk menjadi indikator penting dalam menganalisis peningkatan kebutuhan pangan dari waktu ke waktu.

#### **2.5. Produktivitas Padi dan Ketersediaan Pangan**

Produktivitas padi merupakan salah satu indikator penting dalam mengukur kemampuan suatu daerah dalam memenuhi kebutuhan pangan secara mandiri. Padi sebagai bahan pangan utama memiliki peran strategis dalam menjaga stabilitas konsumsi masyarakat.

Peningkatan produktivitas padi dapat berkontribusi terhadap peningkatan ketersediaan pangan dan mengurangi ketergantungan terhadap pasokan dari luar daerah (Nigus & Dorseywamy, 2021).

Namun, dalam praktiknya, peningkatan produktivitas tidak selalu diikuti dengan peningkatan konsumsi masyarakat secara signifikan. Hal ini disebabkan oleh faktor lain seperti distribusi pangan dan daya beli masyarakat (Pan et al., 2022).

Dengan demikian, produktivitas padi perlu dianalisis bersama dengan faktor lain untuk memahami kondisi ketahanan pangan secara menyeluruh.

## **2.6. Pengaruh Konsumsi Kalori terhadap Kebutuhan Pangan**

Konsumsi kalori merupakan salah satu indikator penting dalam menilai kecukupan gizi dan kebutuhan pangan suatu individu maupun kelompok masyarakat. Kalori menggambarkan jumlah energi yang diperoleh dari makanan dan digunakan untuk menjalankan aktivitas sehari-hari.

Tingkat konsumsi kalori dipengaruhi oleh berbagai faktor, seperti usia, jenis kelamin, tingkat aktivitas, serta kondisi ekonomi. Individu dengan aktivitas tinggi cenderung membutuhkan asupan kalori yang lebih besar dibandingkan dengan individu yang memiliki aktivitas rendah.

Dalam konteks masyarakat, konsumsi kalori sering digunakan sebagai indikator untuk mengukur tingkat kesejahteraan dan ketahanan pangan. Masyarakat dengan tingkat konsumsi kalori yang rendah berpotensi mengalami kekurangan gizi, sedangkan konsumsi kalori yang cukup menunjukkan terpenuhinya kebutuhan energi harian.

Selain itu, konsumsi kalori juga berkaitan erat dengan pola konsumsi pangan. Perubahan dalam konsumsi kalori dapat mencerminkan perubahan pola

makan masyarakat, baik dari segi jumlah maupun jenis makanan yang dikonsumsi.

Dalam penelitian ini, meskipun data yang digunakan berbasis pengeluaran pangan, konsumsi kalori tetap menjadi konsep penting yang mendasari analisis kebutuhan pangan. Hal ini karena pengeluaran terhadap bahan pangan secara tidak langsung mencerminkan tingkat konsumsi energi yang diperoleh masyarakat.

Namun demikian, tidak semua peningkatan pengeluaran pangan secara langsung berbanding lurus dengan peningkatan konsumsi kalori. Hal ini disebabkan oleh perbedaan harga, jenis pangan, serta preferensi konsumsi masyarakat.

Dengan demikian, konsumsi kalori dapat dijadikan sebagai salah satu dasar dalam memahami kebutuhan pangan, meskipun dalam penelitian ini tidak digunakan sebagai variabel utama dalam pemodelan *machine learning*.

## **2.7. Pengaruh Konsumsi Protein terhadap Kebutuhan Pangan**

Konsumsi protein merupakan salah satu indikator penting dalam menilai kualitas konsumsi pangan masyarakat. Protein berperan dalam pertumbuhan, perbaikan jaringan tubuh, serta menjaga fungsi metabolisme. Oleh karena itu, kecukupan konsumsi protein menjadi salah satu aspek penting dalam pemenuhan gizi yang seimbang.

Sumber protein dalam konsumsi masyarakat umumnya berasal dari bahan pangan seperti daging, ikan, telur, susu, serta kacang-kacangan. Variasi konsumsi sumber protein ini sangat dipengaruhi oleh tingkat pendapatan, harga pangan, serta ketersediaan bahan makanan di suatu daerah.

Dalam konteks analisis konsumsi pangan, konsumsi protein sering digunakan sebagai indikator kualitas, sedangkan konsumsi kalori digunakan sebagai indikator kuantitas. Masyarakat dengan konsumsi protein yang cukup cenderung memiliki pola konsumsi yang lebih baik dibandingkan dengan masyarakat yang hanya memenuhi kebutuhan kalori tanpa memperhatikan kualitas gizi.

Selain itu, perubahan pola konsumsi protein juga dapat mencerminkan perubahan tingkat kesejahteraan masyarakat. Peningkatan konsumsi protein hewani, misalnya, sering dikaitkan dengan meningkatnya daya beli masyarakat.

Dalam penelitian ini, meskipun data yang digunakan tidak secara langsung mengukur konsumsi protein, beberapa variabel seperti konsumsi daging, ikan, telur, dan susu dapat dianggap sebagai representasi tidak langsung dari asupan protein masyarakat.

Dengan demikian, konsumsi protein memiliki peran penting dalam analisis kebutuhan pangan, terutama dalam menilai keseimbangan antara kuantitas dan kualitas konsumsi masyarakat.

Namun, keterbatasan data menyebabkan konsumsi protein tidak dijadikan sebagai variabel utama dalam pemodelan. Meskipun begitu, konsep konsumsi protein tetap relevan sebagai dasar dalam memahami kualitas konsumsi pangan masyarakat. Oleh karena itu, penelitian ini menggunakan pendekatan pengeluaran pangan sebagai indikator tidak langsung dari konsumsi masyarakat.

## 2.8. Prediksi

Prediksi merupakan proses memperkirakan nilai masa depan berdasarkan pola data historis. Dalam konteks penelitian ini, prediksi digunakan untuk memperkirakan kebutuhan pangan masyarakat pada periode mendatang dengan memanfaatkan data historis pengeluaran konsumsi pangan.

*Data mining* memiliki beberapa kelompok diantaranya Prediksi. Prediksi merupakan proses memperkirakan nilai atau kejadian di masa depan berdasarkan data yang ada. Hal ini bertujuan untuk menghasilkan perkiraan yang logis atau memiliki probabilitas tinggi tentang apa yang mungkin terjadi di masa depan. Dalam *data mining*, prediksi menggunakan teknik statistik, *machine learning*, atau algoritma prediktif lainnya untuk menemukan pola atau tren dalam data yang dapat dipakai untuk memprediksi nilai atau kejadian di masa depan. (Karim, A, 2024).

## 2.9. Pengeluaran Konsumsi Pangan sebagai Proksi Kebutuhan Pangan

Dalam studi ekonomi pangan, pengeluaran konsumsi pangan rumah tangga kerap digunakan sebagai indikator proksi untuk mengukur kebutuhan dan tingkat konsumsi pangan masyarakat. Pendekatan ini dipilih karena data konsumsi pangan dalam satuan fisik bersifat sensitif dan tidak selalu tersedia untuk publikasi. Oleh sebab itu, pengeluaran per kapita menurut kelompok komoditas pangan menjadi ukuran yang umum diterapkan dalam analisis konsumsi pangan.

Badan Pusat Statistik (BPS) melalui Survei Sosial Ekonomi Nasional (Susenas) menyediakan data konsumsi pangan dalam bentuk rata-rata pengeluaran per kapita per bulan berdasarkan kelompok komoditas. Data tersebut

mencerminkan nilai ekonomi dari konsumsi pangan masyarakat dan dapat digunakan untuk menganalisis perubahan pola konsumsi serta kebutuhan pangan secara agregat. Selain itu, pemanfaatan data pengeluaran konsumsi pangan memungkinkan dilakukan analisis perbandingan antarperiode maupun antarwilayah secara lebih konsisten dan sistematis.

### **2.10. Data Statistik Pangan Badan Pusat Statistik**

Badan Pusat Statistik (BPS) merupakan lembaga pemerintah resmi yang bertanggung jawab atas penyediaan data statistik nasional maupun daerah. Dalam konteks konsumsi pangan, BPS mengumpulkan data melalui Survei Sosial Ekonomi Nasional (Susenas) yang dilaksanakan secara periodik. Data yang diperoleh mencakup berbagai indikator sosial-ekonomi, termasuk pengeluaran konsumsi pangan rumah tangga berdasarkan kelompok komoditas.

Di tingkat kabupaten, BPS menyajikan data dalam bentuk agregat untuk menjaga kerahasiaan identitas responden. Meskipun disajikan secara agregat, data ini memiliki tingkat reliabilitas yang tinggi dan sering digunakan sebagai acuan dalam penelitian akademik maupun penyusunan kebijakan publik. Oleh karena itu, pemanfaatan data BPS Kabupaten Labuhanbatu dalam penelitian ini dianggap relevan dan sesuai dengan standar metodologi ilmiah.



Gambar 2.1 Logo Badan Pusat Statistik

### **2.11. *Machine Learning* dalam Analisis Konsumsi dan Kebutuhan Pangan**

*Machine Learning* merupakan salah satu cabang dari kecerdasan buatan (*artificial intelligence*) yang fokus pada pengembangan algoritma dan model statistik yang memungkinkan komputer belajar dari data serta meningkatkan kinerjanya secara otomatis tanpa pemrograman eksplisit. Menurut Mitchell, *machine learning* didefinisikan sebagai metode yang memungkinkan sistem komputer memanfaatkan pengalaman untuk meningkatkan kinerja dalam melaksanakan suatu tugas tertentu. Dalam konteks analisis konsumsi dan kebutuhan pangan, *machine learning* berperan penting dalam memodelkan hubungan kompleks antara berbagai variabel ekonomi dan sosial yang sulit ditangkap melalui metode statistik tradisional.

Penerapan *machine learning* dalam studi pangan semakin meningkat seiring dengan tersedianya data statistik yang lebih lengkap dan kebutuhan akan prediksi yang lebih akurat. Data konsumsi pangan, yang umumnya berbentuk deret waktu (*time series*), sering kali memiliki pola nonlinier, fluktuasi musiman, serta dipengaruhi oleh faktor eksternal seperti perubahan harga, tingkat pendapatan, dan

pertumbuhan jumlah penduduk. Melalui proses pelatihan model berdasarkan data historis, *machine learning* mampu menangkap pola-pola tersebut secara efektif.

Dalam penelitian konsumsi pangan, *machine learning* tidak hanya digunakan untuk prediksi, tetapi juga untuk analisis tren dan klasifikasi. Analisis tren bertujuan untuk mengidentifikasi arah perubahan konsumsi pangan dari waktu ke waktu, sedangkan klasifikasi digunakan untuk mengelompokkan kondisi konsumsi ke dalam kategori tertentu, misalnya meningkat, menurun, atau stabil. Pendekatan ini memberikan wawasan yang lebih mendalam dibandingkan analisis deskriptif sederhana.

Selain itu, penerapan *machine learning* menawarkan fleksibilitas dalam penyesuaian model sesuai dengan karakteristik dan ketersediaan data, baik data dengan jumlah observasi terbatas maupun data agregat seperti yang dipublikasikan oleh BPS. Dengan demikian, penggunaan *machine learning* sangat relevan untuk menganalisis pengeluaran konsumsi pangan per kapita menurut kelompok komoditas di Kabupaten Labuhanbatu. Namun, keberhasilan metode *machine learning* sangat bergantung pada kualitas dan jumlah data yang digunakan.

Dalam penelitian ini, penerapan *machine learning* difokuskan pada pendekatan *supervised learning* berbasis regresi deret waktu (*time series regression*) karena data memiliki variabel target (label) yang jelas, yaitu pengeluaran konsumsi pangan. Pemilihan pendekatan tersebut didasarkan pada karakteristik data konsumsi pangan yang bersifat tahunan, agregat, serta memiliki jumlah observasi yang terbatas. Oleh karena itu, model yang digunakan menekankan pada kesederhanaan, stabilitas, serta kemudahan interpretasi,

sehingga hasil prediksi yang dihasilkan dapat dipahami dan dimanfaatkan secara langsung dalam konteks perencanaan pangan daerah.

### **2.12. *Data Analytics* dalam Perencanaan Pangan**

*Data Analytics* merupakan proses sistematis yang mencakup pengumpulan, pembersihan, pengolahan, analisis, serta interpretasi data dengan tujuan menghasilkan informasi yang bermakna. Dalam konteks perencanaan pangan, pendekatan ini digunakan untuk memahami pola konsumsi, mengidentifikasi tren jangka panjang, serta mendukung pengambilan keputusan berbasis bukti (*evidence-based decision making*).

Pendekatan *data analytics* memungkinkan penerapan berbagai teknik analisis, mulai dari analisis deskriptif, analisis eksploratif, hingga analisis prediktif. Analisis deskriptif berfungsi untuk memaparkan kondisi konsumsi pangan masyarakat pada periode tertentu, sedangkan analisis eksploratif bertujuan mengungkap pola dan hubungan antarvariabel. Sementara itu, analisis prediktif, yang sering dikombinasikan dengan metode *machine learning*, digunakan untuk memperkirakan kebutuhan pangan di masa depan.

Dalam perencanaan pangan di tingkat daerah, *data analytics* memiliki peran penting dalam menyajikan informasi secara jelas bagi pengambil kebijakan. Visualisasi data, seperti grafik tren dan perbandingan antar kelompok komoditas, membantu pemerintah daerah dalam memahami dinamika konsumsi pangan masyarakat. Dengan demikian, *data analytics* tidak hanya berfungsi sebagai alat analisis, tetapi juga sebagai sarana komunikasi informasi statistik yang efektif.

Penerapan *data analytics* pada data BPS Kabupaten Labuhanbatu memungkinkan analisis yang lebih menyeluruh terhadap pengeluaran konsumsi pangan per kapita. Informasi yang diperoleh dapat digunakan untuk mengidentifikasi kelompok komoditas yang mengalami peningkatan atau penurunan konsumsi, sekaligus menjadi dasar dalam merumuskan kebijakan pangan yang adaptif terhadap perubahan kondisi sosial dan ekonomi masyarakat.

Pendekatan *data analytics* juga memfasilitasi penyajian informasi konsumsi pangan secara sistematis dan mudah dipahami, misalnya melalui tabel ringkasan maupun visualisasi tren. Dalam penelitian ini, *data analytics* digunakan untuk mendukung analisis deskriptif dan eksploratif terhadap pengeluaran konsumsi pangan per kapita berdasarkan kelompok komoditas di Kabupaten Labuhanbatu, sehingga hasil analisis dapat menjadi landasan yang kuat bagi tahap prediksi kebutuhan pangan.

Pendekatan *data analytics* semakin banyak digunakan dalam perencanaan pangan berbasis bukti. Studi terkini menunjukkan bahwa integrasi analisis deskriptif, tren, dan prediktif mampu mendukung perumusan kebijakan pangan yang lebih adaptif terhadap perubahan kondisi sosial ekonomi (Provost & Fawcett, 2023; Phumchusri & Sirimak, 2024).

### **2.13. *Linear Regression***

*Linear Regression* merupakan salah satu metode dalam *machine learning* yang digunakan untuk menggambarkan hubungan antara variabel dependen dengan satu atau lebih variabel independen. Metode ini bertujuan untuk memperkirakan nilai variabel dependen berdasarkan pengaruh variabel

independen yang ada, sehingga dapat dimanfaatkan dalam berbagai kegiatan analisis data maupun proses peramalan (*prediction*) (James et al., 2021).

Menurut James et al. (2021), *linear regression* adalah teknik statistik yang digunakan untuk menganalisis hubungan antara variabel masukan (*input*) dan variabel keluaran (*output*), sekaligus membangun model prediksi yang dapat digunakan untuk memperkirakan nilai pada periode mendatang. Metode ini banyak digunakan dalam analisis data karena memiliki konsep yang relatif sederhana serta menghasilkan model yang mudah dipahami dan diinterpretasikan.

*Linear regression* secara umum dibagi menjadi dua jenis utama, yaitu regresi linear sederhana dan regresi linear berganda. Regresi linear sederhana digunakan ketika hanya terdapat satu variabel independen yang memengaruhi variabel dependen. Sebaliknya, regresi linear berganda digunakan apabila terdapat lebih dari satu variabel independen yang memengaruhi variabel dependen (Kotu & Deshpande, 2019). Dalam penelitian ini digunakan regresi linear berganda karena melibatkan beberapa variabel independen, yaitu jumlah penduduk, konsumsi kalori, dan konsumsi protein, yang digunakan untuk memprediksi total pengeluaran kebutuhan pangan.

Penerapan *linear regression* dalam penelitian memiliki beberapa tujuan utama, di antaranya untuk mengetahui hubungan antara variabel independen dan variabel dependen, melakukan prediksi terhadap nilai variabel dependen pada masa mendatang, serta membantu proses pengambilan keputusan berdasarkan hasil analisis data yang dilakukan (Aggarwal, 2020). Metode ini juga banyak diterapkan dalam penelitian yang berkaitan dengan peramalan kebutuhan atau

tingkat konsumsi karena mampu menunjukkan hubungan antar variabel secara jelas.

*Linear regression* memiliki sejumlah keunggulan, antara lain mudah dipahami, proses analisisnya relatif sederhana, serta mampu menghasilkan nilai prediksi dalam bentuk numerik yang mudah untuk diinterpretasikan. Selain itu, metode ini tidak selalu membutuhkan jumlah data yang sangat besar sehingga masih dapat digunakan pada dataset yang terbatas (James et al., 2021). Hal tersebut menjadikan *linear regression* cocok digunakan dalam penelitian yang menggunakan data tahunan dengan jumlah observasi yang terbatas seperti pada penelitian ini.

Meskipun demikian, *linear regression* juga memiliki beberapa keterbatasan. Metode ini hanya efektif apabila hubungan antara variabel independen dan variabel dependen bersifat linear. Selain itu, metode ini cukup sensitif terhadap keberadaan data pencilan (*outlier*) yang dapat memengaruhi hasil prediksi yang dihasilkan (Kotu & Deshpande, 2019). Oleh karena itu, diperlukan pemilihan variabel yang tepat serta proses pengolahan data yang baik sebelum dilakukan analisis.

Dalam penelitian ini, metode *linear regression* digunakan untuk memprediksi kebutuhan pangan pada periode selanjutnya berdasarkan data historis yang tersedia. Proses pengolahan data serta pembangunan model dilakukan dengan menggunakan perangkat lunak *RapidMiner Studio* sehingga diperoleh hasil prediksi yang dapat dimanfaatkan sebagai dasar analisis dalam memperkirakan kebutuhan pangan pada masa mendatang.

## **2.14. *Naïve Bayes***

*Naïve Bayes* merupakan salah satu metode klasifikasi yang berbasis pada Teorema Bayes dengan asumsi bahwa setiap variabel bersifat independen. Metode ini dikenal memiliki perhitungan yang sederhana dan efisien dalam pengolahan data.

*Naïve Bayes* banyak digunakan dalam berbagai penelitian, termasuk dalam bidang pangan, karena kemampuannya dalam mengolah data dengan cepat dan menghasilkan klasifikasi yang cukup baik (Al Sarwoto et al., 2024).

Selain itu, metode ini juga dapat digunakan untuk memprediksi kebutuhan bahan makanan berdasarkan data historis yang tersedia (Hakim et al., 2025).

Meskipun demikian, tingkat akurasi metode *Naïve Bayes* sangat dipengaruhi oleh jumlah dan kualitas data. Dataset yang terbatas dapat menyebabkan model tidak mampu mengenali pola dengan baik, sehingga menghasilkan akurasi yang rendah.

Oleh karena itu, dalam penggunaannya, perlu diperhatikan pemilihan data dan variabel agar hasil yang diperoleh dapat lebih optimal.

## **2.15. Perangkat Lunak**

### **2.15.1 *RapidMiner***

*RapidMiner* adalah perangkat lunak analisis data yang menyediakan *workflow* visual untuk membangun, melatih, dan mengevaluasi model *machine learning*. *RapidMiner* digunakan dalam penelitian ini karena mendukung implementasi algoritma *Linear Regression* dan *Naïve Bayes* serta menyediakan operator evaluasi kinerja model.

*RapidMiner* merupakan salah satu perangkat lunak yang digunakan untuk analisis data, *data mining*, dan *machine learning*. Aplikasi ini menyediakan berbagai *tools* yang memungkinkan pengguna untuk melakukan pengolahan data, pemodelan, serta evaluasi hasil analisis tanpa harus melakukan pemrograman secara langsung.

Dalam penelitian ini, *RapidMiner* digunakan sebagai alat utama untuk melakukan proses analisis data dan penerapan metode *machine learning*, khususnya algoritma *Naïve Bayes*. Data yang telah diproses sebelumnya menggunakan *Microsoft Excel* kemudian diimpor ke dalam *RapidMiner* untuk dilakukan tahap lanjutan, seperti pemisahan data (*split data*), pelatihan model (*training*), serta pengujian model (*testing*).

*RapidMiner* memiliki berbagai operator yang memudahkan proses analisis, seperti operator *Read Excel* untuk membaca data, *Split Data* untuk membagi dataset menjadi data latih dan data uji, serta *Naïve Bayes* untuk membangun model klasifikasi. Selain itu, *RapidMiner* juga menyediakan operator *Performance* untuk mengevaluasi hasil model berdasarkan tingkat akurasi yang dihasilkan.

Keunggulan *RapidMiner* terletak pada kemudahan penggunaannya melalui antarmuka visual berbasis *drag and drop*, sehingga sangat cocok digunakan dalam penelitian yang berfokus pada analisis data tanpa memerlukan kemampuan pemrograman yang kompleks.

Namun demikian, hasil analisis yang diperoleh dari *RapidMiner* sangat bergantung pada kualitas dan jumlah data yang digunakan. Dataset yang terbatas

dapat mempengaruhi kinerja model yang dihasilkan, sehingga perlu dilakukan pengolahan data yang baik sebelum tahap pemodelan.

Dengan demikian, *RapidMiner* berperan sebagai perangkat lunak utama dalam penelitian ini untuk mengimplementasikan metode *Naïve Bayes* dalam menganalisis dan memprediksi tren konsumsi pangan.



Gambar 2.2 Aplikasi *RapidMiner* sebagai Perangkat Lunak Analisis

### 2.15.2 *Microsoft Excel*

*Microsoft Excel* merupakan salah satu perangkat lunak pengolah data yang dikembangkan oleh *Microsoft* dan termasuk dalam paket *Microsoft Office*. Aplikasi ini digunakan dalam berbagai bidang, termasuk penelitian, karena kemampuannya dalam mengelola, mengolah, dan menganalisis data secara sistematis.

Dalam penelitian ini, *Microsoft Excel* digunakan sebagai alat bantu dalam tahap awal pengolahan data. Data yang diperoleh dari Badan Pusat Statistik (BPS) terlebih dahulu disusun dan dirapikan menggunakan *Microsoft Excel* sebelum dilakukan proses analisis lebih lanjut menggunakan perangkat lunak lain.

Fungsi utama *Microsoft Excel* dalam penelitian ini meliputi proses pembersihan data (*data cleaning*), pengelompokan data, serta perhitungan sederhana seperti penjumlahan total pengeluaran pangan. Selain itu, *Excel* juga

digunakan untuk mengubah format data agar sesuai dengan kebutuhan proses analisis pada *RapidMiner*.

*Microsoft Excel* memiliki keunggulan dalam kemudahan penggunaan serta fitur yang cukup lengkap, seperti fungsi statistik dasar, pengolahan tabel, dan visualisasi sederhana dalam bentuk grafik. Hal ini menjadikan *Excel* sebagai alat yang efektif untuk tahap pra-pemrosesan data.

Meskipun demikian, *Microsoft Excel* memiliki keterbatasan dalam hal analisis data berskala besar dan pemodelan *machine learning*. Oleh karena itu, dalam penelitian ini *Excel* hanya digunakan sebagai alat pendukung, sedangkan proses analisis utama dilakukan menggunakan *RapidMiner* dengan metode *Naïve Bayes*.

Dengan demikian, penggunaan *Microsoft Excel* dalam penelitian ini berperan penting dalam memastikan kualitas data yang digunakan sebelum masuk ke tahap pemodelan dan analisis.



Gambar 2.3 Aplikasi *Microsoft Excel* sebagai Perangkat Lunak Pendukung

## 2.16. Analisis *Time – Series* dalam Studi Konsumsi Pangan

Data konsumsi pangan yang dikumpulkan secara periodik dari waktu ke waktu termasuk dalam kategori data deret waktu (*time series*). Analisis *time series*

bertujuan untuk mengidentifikasi pola historis data, termasuk kecenderungan umum (*trend*), fluktuasi jangka pendek, serta perubahan struktural yang terjadi selama periode pengamatan. Dalam konteks konsumsi pangan, analisis ini penting untuk memahami arah perubahan kebutuhan pangan masyarakat secara berkelanjutan.

Komponen utama dalam data deret waktu meliputi tren (*trend*), musiman (*seasonality*), dan komponen acak (*noise*). Tren menggambarkan arah perubahan konsumsi pangan secara keseluruhan, apakah meningkat, menurun, atau relatif stabil. Meskipun data konsumsi pangan dari Badan Pusat Statistik bersifat tahunan dan tidak secara rinci menangkap efek musiman, analisis tren tetap memberikan gambaran kecenderungan konsumsi pangan dalam jangka menengah dan menjadi dasar bagi perencanaan kebutuhan pangan.

### **2.17. Perbandingan Metode Konvensional dan *Machine Learning***

Metode konvensional dalam analisis konsumsi pangan, seperti analisis deskriptif dan regresi linear sederhana, memiliki keunggulan dari segi kemudahan interpretasi. Namun demikian, metode tersebut seringkali terbatas dalam menangkap pola yang kompleks maupun perubahan nonlinier pada data konsumsi pangan.

Sebaliknya, pendekatan *machine learning* menawarkan fleksibilitas yang lebih tinggi dalam memodelkan hubungan antarvariabel. *Machine learning* memungkinkan pemanfaatan data historis untuk menghasilkan prediksi yang adaptif, meskipun jumlah data relatif terbatas. Dalam penelitian ini, *machine*

*learning* diterapkan dalam bentuk model prediktif berbasis regresi deret waktu, yang disesuaikan dengan karakteristik data Badan Pusat Statistik.

Tabel 2.2 Perbandingan Metode Konvensional dan *Machine Learning*

<b>Aspek</b>	<b>Metode Konvensional</b>	<b><i>Machine learning</i></b>
Pendekatan	Statistik deskriptif dan regresi sederhana	Pembelajaran dari data historis
Kompleksitas Data	Terbatas pada hubungan linier	Mampu menangkap pola tren berupa linier & nonlinier
Kebutuhan Data	Relatif kecil	Fleksibel, disesuaikan data
Akurasi Prediksi	Moderat	Lebih adaptif
Relevansi Penelitian	Analisis kondisi saat ini	Analisis tren dan prediksi

### 2.18. Kelompok Komoditas Pangan yang Dianalisis

Kelompok komoditas pangan yang dianalisis dalam penelitian ini mengacu pada klasifikasi yang ditetapkan oleh Badan Pusat Statistik. Penentuan komoditas didasarkan pada tingkat konsumsi masyarakat, ketersediaan data yang memadai, serta keterkaitannya dengan pemenuhan kebutuhan pangan masyarakat Kabupaten Labuhanbatu.

Tabel 2.3 Kelompok Komoditas Pangan yang Dianalisis

<b>No</b>	<b>Kelompok Komoditas</b>	<b>Deskripsi</b>	<b>Sumber Data</b>
1	Padi – padian	Beras dan produk sejenis	BPS Kab. Labuhanbatu
2	Umbi – umbian	Singkong, ubi, dan sejenisnya	BPS Kab. Labuhanbatu
3	Ikan/Udang/Kerang	Sumber protein hewani	BPS Kab. Labuhanbatu
4	Daging	Daging merah dan	BPS Kab.

		ungags	Labuhanbatu
5	Telur dan Susu	Protein dan kalsium	BPS Kab. Labuhanbatu

### 2.19. Kelebihan dan Keterbatasan *Machine Learning* pada Data Skala Kabupaten

Penerapan *machine learning* pada data berskala terbatas, seperti data konsumsi pangan tahunan di tingkat kabupaten, memiliki keunggulan sekaligus keterbatasan. Salah satu keunggulannya adalah kemampuan model untuk memanfaatkan pola historis data secara optimal dan menghasilkan estimasi prediksi yang lebih sistematis dibandingkan dengan analisis deskriptif semata.

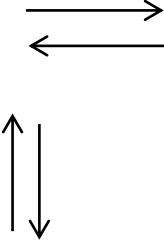


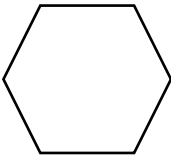
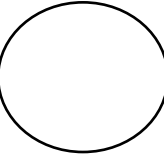

Di sisi lain, keterbatasan jumlah observasi membatasi kompleksitas model yang dapat diterapkan. Oleh karena itu, penelitian ini tidak menggunakan model *machine learning* yang memerlukan volume data besar, seperti *deep learning*, melainkan memilih model yang lebih sederhana namun tetap relevan dan mudah diinterpretasikan. Pendekatan ini dianggap sesuai dengan tujuan penelitian, yang menitikberatkan pada analisis tren serta prediksi awal kebutuhan pangan.

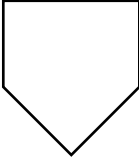
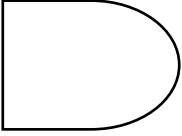

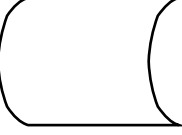
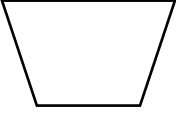
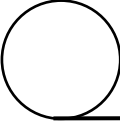
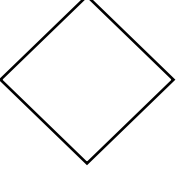



Tabel 2.4 Kelebihan dan Keterbatasan *Machine Learning*

Aspek	Kelebihan	Keterbatasan
Akurasi	Lebih adaptif	Bergantung kualitas data
Fleksibilitas	Menangkap pola tren	Data terbatas
Interpretasi	Cukup jelas	Perlu validasi
Implementasi	Relevan untuk daerah	Tidak cocok <i>deep learning</i>

## 2.20. Flowchart

*Flowchart* atau diagram alir merupakan gambaran visual yang menunjukkan tahapan atau aktivitas dalam suatu proses, program, atau prosedur. Diagram ini menggunakan simbol-simbol yang telah ditentukan secara standar dan dihubungkan dengan anak panah untuk memperlihatkan arah aliran kontrol atau aliran data, sehingga memudahkan pembaca dalam memahami logika serta urutan kerja dari suatu sistem (Sutanti et al., 2020).

Simbol	Nama/Keterangan	Simbol	Nama/Keterangan
	<p><b>Flow Direction Symbol</b></p> <p>Yaitu simbol yang digunakan untuk menghubungkan antara simbol yang satu dengan simbol yang lain. Simbol ini disebut juga <i>connecting line</i>.</p>		<p><b>Symbol Manual Input</b></p> <p>Simbol untuk memasukkan data secara manual <i>online keyboard</i>.</p>
	<p><b>Terminator Symbol</b></p> <p>Yaitu simbol untuk permulaan (<i>start</i>) atau akhir (<i>stop</i>) dari suatu kegiatan.</p>		<p><b>Symbol Preparation</b></p> <p>Simbol untuk mempersiapkan penyimpanan yang akan digunakan sebagai tempat pengolahan di dalam <i>storage</i>.</p>
	<p><b>Connector Symbol</b></p> <p>Yaitu simbol untuk keluar – masuk atau penyambungan proses dalam lembar/halaman yang sama.</p>		<p><b>Symbol Predifine</b></p> <p>Simbol untuk pelaksanaan suatu bagian (sub-program)/prosedur.</p>

	<p><b>Connector Symbol</b> Yaitu simbol untuk keluar – masuk atau penyambungan proses pada lembar/halaman yang berbeda.</p>		<p><b>Symbol Display</b> Simbol yang menyatakan peralatan <i>output</i> yang digunakan yaitu layar, <i>plotter</i>, <i>printer</i> dan sebagainya.</p>
	<p><b>Processing Symbol</b> Simbol yang menunjukkan pengolahan yang dilakukan oleh komputer.</p>		<p><b>Symbol Disk and On-line Storage</b> Simbol yang menyatakan <i>input</i> yang berasal dari <i>disk</i> atau disimpan ke <i>disk</i>.</p>
	<p><b>Simbol Manual Operation</b> Simbol yang menunjukkan pengolahan yang tidak dilakukan oleh komputer.</p>		<p><b>Symbol Magnetik Type Unit</b> Simbol yang menyatakan <i>input</i> berasal dari pita <i>magnetic</i> atau <i>output</i> yang disimpan ke pita <i>magnetic</i>.</p>
	<p><b>Symbol Decision</b> Simbol pemilihan proses berdasarkan kondisi yang ada.</p>		<p><b>Symbol Punch Card</b> Simbol yang menyatakan bahwa <i>input</i> berasal dari kartu atau <i>output</i> ditulis ke kartu.</p>
	<p><b>Symbol Input-Output</b> Simbol yang menyatakan Proses <i>input</i> dan <i>output</i> tanpa tergantung dengan jenis peralatannya.</p>		<p><b>Symbol Document</b> Simbol yang menyatakan <i>input</i> berasal dari dokumen dalam bentuk kertas atau <i>output</i> dicetak ke kertas</p>

Gambar 2.4 *Flowchart*

Fungsi *flowchart* adalah sebagai berikut:

- a. Memvisualkan proses sehingga langkah-langkah menjadi jelas dan terstruktur (Kus Indrani Listyoningrum et al., 2023).
- b. Alat komunikasi antara perancang, *programmer*, dan pemangku kepentingan (dokumentasi) (2023, 2021).
- c. Perancangan dan *debugging* untuk membantu merancang algoritma sebelum *coding* dan menemukan kesalahan logika (Collins et al., 2021b).
- d. Standarisasi Prosedur dalam bisnis / produksi / sistem informasi (Collins et al., 2021a).

### **2.21. Kerangka Pemikiran Penelitian**

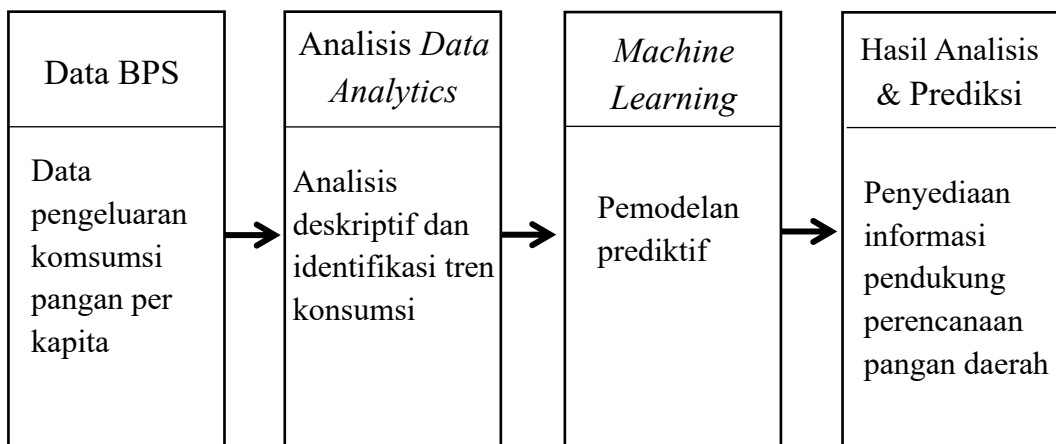
Kerangka pemikiran penelitian ini dirancang untuk menjelaskan tahapan penelitian secara runtut, mulai dari sumber data hingga hasil yang diperoleh. Penelitian memanfaatkan data pengeluaran konsumsi pangan per kapita berdasarkan kelompok komoditas yang berasal dari Badan Pusat Statistik Kabupaten Labuhanbatu sebagai masukan (*input*) utama. Data tersebut digunakan sebagai indikator kebutuhan pangan masyarakat dan dianalisis melalui pendekatan *data analytics* guna mengidentifikasi pola serta kecenderungan tren konsumsi pangan secara sistematis.

Selanjutnya, hasil identifikasi pola dan tren konsumsi pangan tersebut dijadikan landasan dalam penerapan metode *machine learning* untuk membangun model prediksi kebutuhan pangan masyarakat Kabupaten Labuhanbatu pada periode selanjutnya. Proses pemodelan prediktif dilakukan dengan memanfaatkan

keterkaitan historis antarvariabel yang terdapat dalam data konsumsi pangan per kapita.

Keluaran (*output*) penelitian ini berupa informasi mengenai tren konsumsi pangan dan hasil prediksi kebutuhan pangan masyarakat. Informasi tersebut diharapkan mampu memberikan gambaran yang menyeluruh mengenai kondisi serta proyeksi kebutuhan pangan di Kabupaten Labuhanbatu, sehingga dapat dimanfaatkan sebagai dasar dalam perencanaan dan pengambilan kebijakan pangan yang lebih efektif dan berkelanjutan di tingkat daerah.

Secara konseptual, alur kerangka pemikiran penelitian ini dapat digambarkan sebagai berikut:



Gambar 2.5 Kerangka Pemikiran Penelitian

Gambar 2.5 menggambarkan alur penelitian yang dimulai dari pemanfaatan data pengeluaran konsumsi pangan per kapita menurut kelompok komoditas yang bersumber dari Badan Pusat Statistik (BPS) Kabupaten Labuhanbatu, yang selanjutnya dianalisis menggunakan pendekatan *data analytics* untuk melakukan analisis deskriptif serta mengidentifikasi pola dan tren konsumsi pangan masyarakat, kemudian hasil analisis tersebut digunakan sebagai dasar dalam

penerapan metode *machine learning* untuk membangun model prediktif kebutuhan pangan pada periode mendatang, sehingga menghasilkan keluaran berupa informasi prediksi kebutuhan pangan yang diharapkan dapat mendukung perencanaan dan pengambilan kebijakan pangan di tingkat daerah, dengan uraian teknis setiap tahapan penelitian dijelaskan lebih lanjut pada Bab III.