

BAB I

PENDAHULUAN

1.1 Latar Belakang

Industri makanan pada sektor Usaha Mikro, Kecil, dan Menengah (UMKM), khususnya produk roti, menghadapi tantangan besar terkait umur simpan produk yang pendek dan fluktuasi permintaan pasar. Ketidakakuratan dalam memprediksi permintaan sering kali memicu kerugian ganda: *overstock* yang berujung pada pembuangan produk (limbah makanan), atau kehabisan stok yang menghilangkan potensi pendapatan. Permasalahan mendesak ini tengah dialami oleh Toko HN Bakery, di mana sistem perencanaan produksi harian masih mengandalkan perkiraan dan belum selaras dengan dinamika permintaan konsumen. Jika kondisi ini dibiarkan tanpa intervensi sistematis, HN Bakery berisiko mengalami kebocoran finansial yang berkelanjutan dan inefisiensi operasional di tengah iklim persaingan bisnis yang semakin ketat (Harris, wawancara, 18 Juli 2026).

Tantangan utama dalam menyelaraskan produksi adalah kompleksitas pola penjualan ritel yang dipengaruhi oleh multi-faktor, seperti variasi produk, harga, promosi, hingga pola waktu harian dan musiman. Saat ini, metode prediksi konvensional berbasis intuisi sangat terbatas dalam mengolah data historis yang kompleks sehingga rentan memicu alokasi sumber daya yang salah sasaran. Mengingat hubungan antar variabel penjualan sering kali bersifat non-linear, pendekatan tradisional tidak lagi memadai. Oleh karena itu, diperlukan intervensi *machine learning* yang terbukti lebih adaptif dalam memodelkan pola deret waktu historis ritel, sebagaimana ditunjukkan oleh efektivitas *Random Forest* (Efendi *et al.*, 2024) dan stabilitas *Support Vector Regression* (Zakir *et al.*, 2025) pada penelitian-penelitian sebelumnya.

Untuk menjawab kebutuhan sistem peramalan yang presisi dan andal, algoritma *Gradient Boosted Decision Tree* (GBDT) hadir sebagai pendekatan komputasi yang sangat sesuai untuk memodelkan karakteristik data tabular UMKM. Metode yang diperkenalkan oleh Jerome H. Friedman ini bekerja dengan membangun pohon keputusan secara sekuensial guna meminimalkan kesalahan prediksi melalui optimasi *gradient descent*. Keunggulan utama GBDT terletak pada kapabilitasnya menangkap pola interaksi non-linear yang kompleks dari berbagai variabel penjualan tanpa memerlukan transformasi matematis yang rumit. Karakteristik ini menjadikan GBDT sebagai metode

yang sangat representatif untuk diterapkan langsung dalam menggali wawasan dari data transaksi operasional.

Sebagai langkah konkret untuk menyelesaikan permasalahan pada HN Bakery, penelitian ini mengimplementasikan model prediksi GBDT menggunakan data transaksi periode pertengahan 2022 hingga 2025 yang diagregasi secara mingguan. Model ini dioptimalkan menggunakan rekayasa fitur kalender (*calendar engineering*) yang memisahkan variabel *weekend* dan *weekday* guna merespons sensitivitas penjualan roti terhadap hari libur. Keandalan model akan divalidasi melalui proses *backtesting* dan dievaluasi menggunakan metrik *Mean Absolute Error* (MAE) sebelum digunakan untuk memproyeksikan penjualan tahun 2026. Berfokus murni pada implementasi algoritma GBDT tanpa perbandingan model lain, penelitian ini juga melakukan analisis *feature importance* untuk mengidentifikasi variabel yang paling dominan, sehingga data transaksi HN Bakery dapat ditransformasikan menjadi kebijakan strategis yang solutif.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, maka rumusan masalah dalam penelitian ini adalah sebagai berikut :

1. Bagaimana membangun model prediksi penjualan produk roti menggunakan algoritma Gradient Boosting Decision Tree (GBDT) berdasarkan data historis untuk menghasilkan proyeksi periode berikutnya?
2. Faktor-faktor apa saja yang paling dominan memengaruhi penjualan produk roti berdasarkan analisis feature importance pada model GBDT?
3. Bagaimana mengonversi hasil prediksi numerik ke dalam kategori operasional menggunakan pendekatan kuartil statistik sebagai dasar sistem pendukung keputusan?
4. Seberapa baik performa model GBDT dalam memprediksi penjualan produk roti berdasarkan evaluasi menggunakan Mean Absolute Error (MAE)?

1.3 Tujuan Penelitian

Terdapat beberapa tujuan dari penelitian ini sebagai berikut :

1. Membangun model prediksi penjualan produk roti menggunakan algoritma Gradient Boosting Decision Tree (GBDT) berdasarkan data historis penjualan.
2. Mengidentifikasi faktor-faktor dominan yang memengaruhi penjualan melalui analisis feature importance dari model yang dibangun.
3. Mengembangkan mekanisme klasifikasi hasil prediksi ke dalam kategori “Rendah”, “Sedang”, dan “Tinggi” menggunakan pendekatan kuartil statistik.

4. Mengevaluasi performa model prediksi menggunakan metrik Mean Absolute Error (MAE) untuk mengukur tingkat akurasi prediksi.

1.4 Batasan Penelitian

Agar penelitian ini lebih terarah dan sesuai dengan tujuan yang ingin dicapai, maka batasan masalah dalam penelitian ini adalah sebagai berikut :

1. Objek penelitian terbatas pada Toko HN Bakery sebagai unit usaha skala UMKM.
2. Data yang digunakan merupakan data historis transaksi harian periode juli tahun 2022 – 2025 bulan desember.
3. Variabel dependen adalah jumlah roti terjual, sedangkan variabel independen meliputi jenis roti, harga, kapasitas stok harian, status promosi (Label Mapping), serta ekstraksi elemen waktu kalender (weekend dan weekday).
4. Pembagian data dilakukan menggunakan pendekatan Time-Series Split dengan proporsi $\pm 80\%$ data latih dan $\pm 20\%$ data uji.
5. Metode utama yang digunakan adalah Gradient Boosted Decision Tree (GBDT) mencakup tahap preprocessing, cleaning data, serta penentuan ambang batas kategori operasional menggunakan kuartil statistik.
6. Evaluasi performa difokuskan pada metrik Mean Absolute Error (MAE) yang dikalkulasi secara manual menggunakan rumus matematis di dalam sistem.
7. Sistem yang dibangun berupa dashboard berbasis web menggunakan Flask dan tidak terintegrasi dengan sistem kasir (Point of Sale).
8. Output dari mekanisme klasifikasi berbasis kuartil statistik (Q1, Q3) mencakup pelabelan kategori (Rendah/Sedang/Tinggi) dan rekomendasi tindak lanjut operasional secara otomatis.

1.5 Manfaat

Adapun manfaat yang dapat diperoleh dari penelitian ini, diantaranya sebagai berikut :

1. Memberikan dasar pengambilan keputusan berbasis data bagi pemilik HN Bakery dalam perencanaan produksi dan pengelolaan stok.
2. Mengurangi risiko overstock dan stockout melalui estimasi permintaan dan klasifikasi penjualan berbasis kuartil.
3. Memberikan kontribusi ilmiah dalam pengembangan model prediksi penjualan menggunakan GBDT dan analisis feature importance pada sektor UMKM.
4. Menjadi referensi bagi penelitian selanjutnya dalam pengembangan model prediksi yang terintegrasi dengan klasifikasi operasional dalam sistem pendukung keputusan.

1.6 Sistematika Penulisan

Untuk mempermudah memahami pembahasan pada penulisan skripsi ini, maka sistematika penulisan diperoleh sebagai berikut :

BAB I PENDAHULUAN

Bab ini berisi uraian mengenai latar belakang penelitian, rumusan masalah, batasan penelitian, tujuan penelitian, manfaat penelitian, serta sistematika penulisan yang digunakan dalam penyusunan laporan. Pada bab ini dijelaskan permasalahan yang terjadi pada HN Bakery terkait perencanaan produksi dan prediksi penjualan serta alasan penggunaan metode Gradient Boosted Decision Tree (GBDT) sebagai solusi penelitian.

BAB II TINJAUAN PUSTAKA

Bab ini membahas teori-teori yang mendukung penelitian, meliputi penelitian terdahulu, konsep machine learning, sistem pendukung keputusan, prediksi penjualan, metode Gradient Boosted Decision Tree (GBDT), Feature Importance, Mean Absolute Error (MAE), kuartil statistik, serta teknologi yang digunakan dalam pengembangan sistem seperti Python, Flask, dan MySQL.

BAB III ANALISIS DAN PERANCANGAN SISTEM

Bab ini menjelaskan tahapan analisis kebutuhan dan perancangan sistem yang dibangun. Pembahasan meliputi analisis kebutuhan sistem, arsitektur sistem, perancangan basis data, use case diagram, flowchart sistem, struktur menu, serta perancangan antarmuka pengguna yang akan digunakan pada aplikasi prediksi penjualan HN Bakery.

BAB IV IMPLEMENTASI DAN HASIL

Bab ini menjelaskan proses implementasi sistem yang telah dirancang, meliputi implementasi database, implementasi antarmuka, proses preprocessing data, pembentukan model Gradient Boosted Decision Tree (GBDT), perhitungan manual metode, analisis feature importance, evaluasi menggunakan Mean Absolute Error (MAE), proses backtesting, hasil prediksi penjualan tahun 2026, pengujian website menggunakan metode Black Box Testing, serta pengujian hyperparameter tuning untuk memperoleh parameter terbaik pada model.

BAB V PENUTUP

Bab ini berisi kesimpulan yang diperoleh berdasarkan hasil penelitian dan implementasi sistem yang telah dilakukan. Bab ini juga memuat saran yang dapat digunakan sebagai bahan pengembangan penelitian maupun pengembangan sistem pada masa mendatang agar memperoleh hasil yang lebih optimal.