

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Harga pangan merupakan indikator ekonomi penting yang mencerminkan keseimbangan pasokan dan permintaan komoditas dengan stabilitas harga yang krusial bagi kesejahteraan masyarakat terutama untuk komoditas sembako sebagai kebutuhan dasar sehari-hari [1][2]. Tiga komoditas sembako yang paling berpengaruh adalah Beras, minyak goreng, dan telur yang paling berkontribusi signifikan terhadap inflasi dengan bobot BPS 2012 sebesar 4,19%, 1,22%, dan 0,92% [1-6]. Namun, ketidakstabilan harga ketiga komoditas ini telah menjadi masalah serius di Provinsi Sumatera Utara.

Di Sumatera Utara, harga sembako mencapai level kritis naik-turun drastis dan tak terkendali, mengancam 15,1 juta penduduk. Inflasi awal 2023 mencapai 4,80% (nasional 3,2%) [7], dengan cabai merah melonjak 200-300% dalam sebulan di sembilan kabupaten/kota [6]. Beras (4,19%), minyak goreng (1,22%), dan telur (0,92%) menyumbang 6,33% inflasi [3]. Dampaknya: 3,8 juta keluarga miskin (25% populasi) yang mengalokasikan 50-60% pendapatan untuk pangan sangat terpukul, kerugian ekonomi Rp 2,3 triliun/kuartal [1][2]. Dari anggaran Rp 1,5 triliun/tahun, Rp 600 miliar (40%) terbangun karena operasi pasar terlambat akibat prediksi tidak akurat [4]. Karena itu, prediksi harga yang akurat sangat *urgent* untuk mencegah pemborosan anggaran dan melindungi masyarakat dari lonjakan harga mendadak [1-4].

Untuk mengatasi fluktuasi ekstrem harga sembako di Sumatera Utara yang kerap melonjak hingga 200-300% per bulan [6], diperlukan model sistem *early warning* berbasis *machine learning*. Model ini sangat krusial untuk menggantikan intervensi reaktif pemerintah yang selama ini sering terlambat dan tidak efektif. Dengan kemampuannya mengolah data historis untuk memprediksi harga secara akurat 7-30 hari ke depan [8-12], model ini mengubah pendekatan penanganan menjadi proaktif. Beberapa penelitian sejenis telah membuktikan efektivitasnya, di antaranya prediksi harga sembako di DKI Jakarta menggunakan Regresi Linear Berganda dengan kontribusi prediktor 84,2% [14], prediksi harga beras di Kota Medan menggunakan Regresi Linear Sederhana dengan RMSE 0,559 [13], serta penerapan XGBoost untuk prediksi harga cabai rawit di DKI Jakarta yang menghasilkan  $R^2$  99% pada data latih dan 92% pada data uji [12]. Namun, metode-metode tersebut masih terbatas dalam menangani pola non-linear dan variabel eksternal secara

bersamaan, sehingga dibutuhkan pendekatan yang lebih robust untuk kondisi fluktuasi harga di Sumatera Utara. Secara praktis, penerapan sistem peringatan dini berbasis *machine learning* yang lebih optimal bermanfaat untuk mencegah lonjakan harga sejak awal, menghemat Rp 600 miliar anggaran operasi pasar yang tidak efisien [4], menekan kerugian ekonomi hingga Rp 2,3 triliun per kuartal, melindungi 3,8 juta keluarga miskin [1][2], serta mengoptimalkan manajemen stok pangan di 33 kabupaten/kota.

Salah satu algoritma yang menunjukkan performa unggul dalam melakukan prediksi adalah *Extreme Gradient Boosting* (XGBoost). XGBoost mampu mengolah data kompleks dengan banyak fitur, menangani hubungan non-linear, dan dilengkapi regularisasi untuk mencegah *overfitting*. XGBoost secara konsisten mengungguli model lain dengan  $R^2$  hingga 0,82 dan MAPE 5–10%, mendemonstrasikan kemampuannya menangkap hubungan non-linear kompleks dan memperhitungkan faktor eksternal seperti inflasi dan musiman [10][11]. Berbeda dengan metode *time series* murni, XGBoost dapat mengintegrasikan variabel eksternal seperti cuaca, volume pasokan, harga komoditas terkait, dan indikator ekonomi untuk prediksi yang lebih akurat. Penelitian pada cabai rawit menunjukkan XGBoost mencapai  $R^2$  99% pada data latih, 92% pada data uji, dan *cross-validation* 96%, memperkuat keandalan dan ketangguhannya [12].

Agar hasil prediksi ini dapat digunakan secara praktis oleh pemerintah daerah, penelitian ini juga membangun aplikasi web sederhana berbasis Python dengan *framework* Streamlit yang dinamakan SIPAHANG atau Sistem Prediksi Harga Pangan. Aplikasi ini mengintegrasikan model XGBoost ke dalam antarmuka interaktif yang menyajikan visualisasi harga dan proyeksi masa depan. Melalui sistem SIPAHANG, pemangku kebijakan dapat lebih mudah memantau risiko lonjakan harga serta mengambil keputusan stabilisasi pangan secara proaktif berdasarkan data yang akurat.

Berdasarkan uraian tersebut, penelitian ini mengangkat judul Prediksi Harga Penjualan Sembako di Wilayah Sumatera Utara Menggunakan Algoritma XGBoost untuk menghasilkan model prediksi yang dapat membantu pemangku kebijakan dalam mengantisipasi fluktuasi harga pangan.

## 1.2 Rumusan Masalah

Fluktuasi harga sembako di Sumatera Utara perlu diantisipasi secara dini untuk mengurangi kerugian ekonomi sehingga diperlukan model prediksi berbasis *machine learning* untuk mendukung intervensi yang lebih tepat waktu.

### 1.3 Tujuan

Tujuan dari penelitian ini adalah untuk menghasilkan model yang dapat memprediksi harga sembako di wilayah Sumatera Utara pada rentang tertentu berdasarkan data yang telah ada, dengan menggunakan metode *Extreme Gradient Boosting* (XGBoost).

### 1.4 Manfaat

Adapun manfaat yang diharapkan dari penelitian ini dapat membantu pihak terkait, dan pemangku kebijakan, dalam memperoleh informasi prediksi harga komoditas pangan yang lebih akurat sebagai dasar pendukung pengambilan keputusan dalam upaya pengendalian dan stabilisasi harga pangan secara cepat.

### 1.5 Ruang Lingkup

Ruang lingkup dari penelitian ini adalah sebagai berikut:

1. Data yang digunakan adalah harga mingguan komoditas sembako meliputi beras, minyak goreng, dan telur di wilayah Sumatera Utara periode 2020-2024 yang menghasilkan sekitar 783 observasi. Ketiga komoditas ini dipilih karena berkontribusi signifikan terhadap inflasi dan merupakan penyumbang utama inflasi *volatile food* dengan ketersediaan data historis lengkap pada situs <https://www.bi.go.id/hargapangan/TabelHarga/ProdusenKomoditas> yang dikelola oleh Bank Indonesia melalui Pusat Informasi Harga Pangan Strategis atau PIHPS [1-6].
2. Wilayah penelitian difokuskan pada Provinsi Sumatera Utara dengan periode data tahun 2020–2024, dengan data 2020-2023 untuk pelatihan dan 2024 untuk pengujian.
3. Variabel penelitian terdiri dari variabel target berupa harga komoditas sembako dan variabel input berupa data historis harga serta variabel ekonomi makro yaitu inflasi (BPS) dan nilai tukar rupiah (yfinance). Pemilihan variabel ini didasarkan pada ketersediaan data historis yang lengkap dan kredibel sehingga memungkinkan proses pelatihan model yang lebih optimal.
4. Algoritma yang digunakan adalah *Extreme Gradient Boosting* (XGBRegressor) dengan pendekatan *supervised learning*. Parameter learning XGBoost yang digunakan meliputi *learning rate* (0.01–0.3), *max depth* (3–7), *n\_estimators* (300–800), *subsample* (0.7–0.9), *colsample\_bytree* (0.7–0.9), serta parameter regularisasi *lambda/L2* (0.5–1.5) dan *alpha/L1* (0–0.5). Optimasi parameter dilakukan melalui *grid search*

dengan *Time Series Split cross-validation* ( $k=5$ ) untuk meminimalkan nilai MAE sebagai metrik evaluasi utama.

5. Evaluasi kinerja model menggunakan metrik *Mean Absolute Error* (MAE) sebagai ukuran akurasi utama serta didukung oleh *Mean Squared Error* (MSE) dan *Mean Absolute Percentage Error* (MAPE) untuk memberikan gambaran kesalahan prediksi yang lebih menyeluruh.
6. Output penelitian diimplementasikan dalam bentuk aplikasi web sederhana yang dibangun menggunakan bahasa pemrograman Python. Berfungsi untuk memvisualisasikan hasil prediksi harga sembako secara interaktif sehingga dapat memberikan gambaran yang jelas mengenai potensi fluktuasi harga kepada pengguna.

