

BAB I

PENDAHULUAN

1.1. Latar Belakang

Sektor pertanian memiliki peran yang sangat penting dalam perekonomian Indonesia, termasuk dalam bidang kebutuhan pangan. Salah satu sektor pertanian yang cukup luas berada di Provinsi Jawa Timur yang biasanya juga dikenal dengan luasnya lahan pertanian serta keberagaman komoditas unggulan yang dihasilkannya[1]. Dari banyaknya komoditas unggulan yang dihasilkan, Jawa Timur memiliki 38 Kabupaten/Kota yang menjadi dasar penghasil komoditas unggulan. Komoditas Unggulan di Jawa Timur adalah Jagung, Kedelai, Ubi Jalar, Ubi kayu, Kacang Tanah, dan Kacang Hijau. Sebagai salah satu daerah penghasil utama produk pertanian, Jawa Timur memiliki kontribusi besar dalam pemenuhan kebutuhan pangan nasional. Luas lahan pertanian di Kabupaten/Kota Se-Provinsi Jawa Timur mencapai 1.214.909 hektare pada tahun 2021, setara dengan 25,41% dari total luas wilayah Jawa Timur[2]. Keanekaragaman komoditas unggulan yang dihasilkan menjadikan provinsi ini sebagai salah satu penyokong utama sektor pertanian nasional.

Dalam beberapa tahun terakhir, isu ketahanan pangan semakin memperoleh perhatian sebagai bagian dari prioritas pembangunan nasional. Hal ini tercermin dalam kebijakan strategis seperti program *Makan Bergizi Gratis (MBG)*, yang menandai pergeseran fokus anggaran nasional dari sektor pendidikan dan kesehatan menuju sektor pangan dalam rangka mewujudkan visi Indonesia Emas 2045[3]. Di tingkat regional, Provinsi Jawa Timur memegang peranan penting dalam mendukung ketahanan pangan nasional melalui enam subsektor utama dalam sektor pertanian, yaitu tanaman pangan, hortikultura, perkebunan, peternakan, kehutanan, dan perikanan. Berdasarkan data Badan Pusat Statistik Provinsi Jawa Timur (2024), jumlah Rumah Tangga Usaha Pertanian (RTUP) di provinsi ini tercatat sebanyak 5.531.730 rumah tangga. Sementara itu, Sensus Pertanian 2023 mencatat bahwa terdapat 5.678.945 unit usaha pertanian, yang terdiri dari 5.676.717 usaha perorangan (UTP), 637 badan usaha berbadan hukum (UPB), dan 1.591 usaha lainnya (UTL) [4].

Salah satu komoditas unggulan di Jawa Timur adalah Jagung. Komoditas jagung memiliki peranan strategis dalam sistem ketahanan pangan nasional karena digunakan tidak hanya sebagai makana pokok utama. Jagung juga menjadi sumber pangan pokok kedua setelah beras dan memiliki peran penting dalam sektor peternakan sebagai bahan baku pakan serta dalam industri pengolahan makanan dan bioenergi. Karena peran gandanya tersebut, kestabilan harga jagung menjadi sangat penting untuk menjamin kesejahteraan petani, menghindari keresahan konsumen, dan menjaga keberlanjutan industri terkait[5]. Harga yang stabil memungkinkan petani mendapatkan keuntungan yang layak, serta memastikan konsumen dapat memperoleh jagung dengan harga terjangkau.

Kabupaten Sampang di Pulau Madura memiliki potensi pertanian yang besar, terutama dalam produksi jagung sebagai komoditas strategis yang digunakan untuk pakan ternak, industri makanan olahan, dan bioenergi. Permintaan jagung yang tinggi menjadikannya kunci dalam mendukung ketahanan pangan nasional dan regional, sehingga menjaga stabilitas harganya sangat penting bagi kesejahteraan petani dan keberlanjutan industri terkait [6]. Namun, fluktuasi harga jagung masih menjadi tantangan akibat pengaruh faktor musim, biaya distribusi, permintaan pasar, serta kebijakan perdagangan seperti impor dan subsidi. Saat panen raya harga sering turun, sedangkan pada musim paceklik harga melonjak, yang berdampak pada pendapatan petani, daya beli masyarakat, dan stabilitas pangan nasional [5][7].

Berdasarkan data SISKAPERBAKO Jawa Timur (2022–2024), harga jagung konsumen di Kabupaten Sampang menunjukkan tren fluktuatif, dari sekitar Rp 9.500/kg pada awal 2022 turun menjadi Rp 8.500/kg pada akhir 2024. Perubahan ini dipengaruhi oleh faktor musim, distribusi antarpulau, dan curah hujan yang tidak merata, sehingga mengindikasikan ketidakstabilan harga yang dapat berdampak pada pendapatan petani dan daya beli masyarakat. Oleh karena itu, diperlukan sistem prediksi berbasis data untuk mendukung kebijakan stabilisasi harga pangan daerah. Pendekatan berbasis teknologi seperti *machine learning*, khususnya *algoritma Random Forest Regression* (RFR), mampu menganalisis pola non-linear dan kompleks pada data historis harga jagung, sehingga dapat memetakan serta memprediksi tren harga dengan lebih akurat[8].

Random Forest Regression (RFR) adalah metode yang mengandalkan ensemble learning dengan membangun banyak pohon keputusan (*decision trees*) dan menggabungkan hasilnya untuk menghasilkan prediksi yang lebih akurat. Setiap pohon dilatih menggunakan subset data yang diambil secara acak, memberikan perspektif berbeda terhadap data[9]. Proses agregasi hasil dari semua pohon membantu mengurangi variabilitas dan meningkatkan stabilitas prediksi. Keunggulan utama metode ini adalah kemampuannya mengatasi data yang kompleks dan bersifat non-linear, serta efektivitasnya dalam mengidentifikasi variabel-variabel penting yang berkontribusi terhadap hasil prediksi. Dengan demikian, penerapan *Random Forest Regression* (RFR) dalam memprediksi kebutuhan pupuk diharapkan dapat meningkatkan efisiensi distribusi, produktivitas pertanian, dan mendukung keputusan berbasis data yang lebih baik [10].

Penelitian sebelumnya membuktikan bahwa *Random Forest Regression* (RFR) memiliki performa unggul dalam prediksi agrikultur, salah satunya pada kasus harga beras di Jawa Barat dengan akurasi mencapai 98,69%, lebih tinggi dibandingkan regresi. Hal ini menunjukkan potensi metode tersebut dalam konteks prediksi agribisnis[11]. Selain itu, penelitian di Jawa Timur menerapkan metode *Random Forest* dan *Geographically Weighted Random Forest* untuk menganalisis faktor-faktor stunting. Penelitian ini membuktikan bahwa pendekatan *Random Forest* cukup andal dalam memodelkan fenomena spasial berbasis data kuantitatif dengan hasil prediksi yang akurat.

Kemudian, untuk menghasilkan sistem peramalan yang lebih optimal, algoritma *Random Forest Regression* (RFR) dipandang sebagai salah satu metode *machine learning* yang memiliki potensi tinggi. Algoritma ini memiliki kemampuan dalam mengolah data dengan hubungan non-linier serta mengurangi risiko *overfitting* melalui teknik *ensemble* berbasis banyak pohon keputusan. Selain itu, RFR juga memungkinkan dilakukannya analisis terhadap tingkat kepentingan variabel (*feature importance*), yang berfungsi untuk mengidentifikasi fitur-fitur input yang memberikan kontribusi terbesar terhadap hasil prediksi [12].

Selanjutnya, penelitian ini mengadopsi metodologi CRISP-DM (*Cross-Industry Standard Process for Data Mining*) sebagai kerangka kerja yang sistematis untuk mengarahkan proses analisis data[13]. Metodologi ini memungkinkan peneliti

untuk menjalankan setiap tahap analisis dengan lebih terstruktur, mulai dari pemahaman bisnis hingga evaluasi model. Dalam upaya untuk mengevaluasi kinerja model yang dikembangkan, penelitian ini menggunakan empat metrik evaluasi yang umum digunakan, yaitu *Mean Absolute Error (MAE)*, *Mean Squared Error (MSE)*, *Root Mean Squared Error (RMSE)* dan *coefficient of determination (R^2)*[14].

Sebagai upaya untuk memastikan keandalan dan keakuratan model prediksi, penelitian ini juga akan membandingkan kinerja metode *Random Forest Regression (RFR)* dengan salah satu metode peramalan statistik klasik, yaitu *Double Exponential Smoothing*. Metode *Double Exponential Smoothing* dikenal efektif untuk meramalkan data deret waktu yang memiliki pola tren, karena menggunakan dua parameter peluruhan (level dan tren) untuk menyesuaikan perubahan nilai dari waktu ke waktu. Perbandingan ini dilakukan untuk menilai sejauh mana algoritma berbasis machine learning seperti *Random Forest* dapat memberikan keunggulan dibandingkan metode statistik tradisional dalam memprediksi harga jagung konsumen[15].

Selain itu, untuk membangun sistem peramalan yang lebih akurat, metode *Random Forest Regression (RFR)* menjadi salah satu pendekatan *machine learning* yang menjanjikan. RFR mampu menangani data *non-linear*, meminimalkan *overfitting* melalui mekanisme *ensemble decision trees*, serta menghasilkan model yang stabil meskipun terdapat variasi pada data input[16]. Keunggulan lainnya adalah kemampuannya dalam melakukan analisis *feature importance*, yaitu mengidentifikasi variabel-variabel input yang paling berpengaruh terhadap nilai prediksi[17].

Dengan adanya penelitian berjudul **“Prediksi dan Analisis Tren Harga Jagung Pada Pasar Tradisional di Kabupaten Sampang Menggunakan Algoritma *Random Forest Regression*”**, dapat memberikan kontribusi nyata dalam menghadapi permasalahan fluktuasi harga jagung di daerah. Penelitian ini tidak hanya berfokus pada pembangunan model prediksi harga yang akurat, tetapi juga pada upaya menghadirkan inovasi teknologi berbasis *machine learning* yang dapat dimanfaatkan sebagai alat bantu pengambilan keputusan di sektor pertanian.

Berdasarkan uraian tersebut, dapat diidentifikasi bahwa permasalahan utama dalam penelitian ini terletak pada ketidakstabilan harga jagung yang masih tinggi

dan sulit diprediksi, sehingga menimbulkan ketidakpastian bagi petani maupun konsumen dalam menentukan strategi produksi dan konsumsi. Selain itu, hingga saat ini belum tersedia sistem prediksi harga jagung yang berbasis data historis di Kabupaten Sampang, sehingga kebijakan dan keputusan terkait harga masih banyak bergantung pada perkiraan musiman yang bersifat subjektif. Di sisi lain, penerapan metode analisis modern berbasis *machine learning* dalam kebijakan harga pangan daerah juga masih sangat terbatas, padahal pendekatan tersebut memiliki potensi besar untuk menghasilkan prediksi yang lebih akurat dan mendukung pengambilan keputusan yang berbasis data (*data-driven decision making*).

Dengan mempertimbangkan permasalahan tersebut, penelitian ini berupaya untuk mengembangkan model prediksi harga jagung menggunakan algoritma *Random Forest Regression* (RFR) sebagai solusi berbasis teknologi yang mampu memprediksi fluktuasi harga secara lebih tepat. Melalui penerapan model ini, diharapkan dapat diperoleh sistem prediksi harga yang akurat dan aplikatif untuk membantu pemerintah daerah, pelaku pasar, serta petani dalam merencanakan produksi dan distribusi komoditas jagung secara lebih efektif.

1.2. Perumusan Masalah

Harga jagung konsumen di Kabupaten Sampang masih menunjukkan fluktuasi yang cukup signifikan dari tahun ke tahun. Kondisi tersebut menimbulkan ketidakpastian bagi petani, pedagang, maupun pengambil kebijakan dalam menentukan strategi produksi, distribusi, serta pengendalian harga di pasar. Metode prediksi konvensional yang berbasis model statistik sederhana belum mampu menangkap pola data yang bersifat non-linear dan dinamis secara optimal, sehingga hasil prediksi sering kali kurang akurat. Oleh karena itu, diperlukan pendekatan berbasis *machine learning* yang lebih adaptif dan mampu mengenali pola kompleks dalam data historis harga jagung. Berdasarkan permasalahan yang telah dijelaskan, penelitian ini menerapkan metode *machine learning*.

1.2.1. Metode Usulan

Untuk menjawab permasalahan fluktuasi harga jagung, penelitian ini mengusulkan penggunaan *algoritma Random Forest Regression* (RFR) sebagai metode utama. RFR dipilih karena:

1. Mampu menangkap pola data yang kompleks dan non-linear.

2. Mengurangi risiko overfitting melalui teknik ensemble learning.
3. Memberikan analisis *feature importance* sehingga dapat mengidentifikasi variabel yang paling berpengaruh terhadap harga.

Sebagai pembandingan, digunakan metode *Double Exponential Smoothing* (DES), yang mewakili pendekatan peramalan deret waktu klasik. Perbandingan kedua metode ini bertujuan untuk menilai sejauh mana algoritma berbasis machine learning dapat memberikan prediksi yang lebih akurat dibandingkan metode statistik tradisional

1.2.2. Pertanyaan Penelitian

Berdasarkan permasalahan dan metode yang diterapkan, penelitian ini merumuskan pertanyaan penelitian sebagai berikut:

1. Apakah penerapan *algoritma Random Forest Regression* (RFR) efektif dalam membangun model prediksi harga jagung konsumen di Kabupaten Sampang berdasarkan data historis periode 2022–2024?
2. Berapa nilai performa yang diperoleh oleh model *Random Forest Regression* (RFR) dibandingkan dengan metode *Double Exponential Smoothing* (DES) berdasarkan metrik evaluasi MAE, MSE, RMSE, dan R^2 ?
3. Faktor-faktor apa saja yang berperan paling dominan terhadap fluktuasi harga jagung konsumen menurut hasil analisis *feature importance* dari model *Random Forest Regression*?

1.3. Tujuan dan Manfaat

1.3.1. Tujuan Penelitian

1. Membangun model prediksi harga jagung konsumen di Kabupaten Sampang periode 2022–2024 menggunakan algoritma *Random Forest Regression* (RFR) berbasis pendekatan CRISP-DM.
2. Mengevaluasi kinerja algoritma RFR dengan membandingkannya terhadap metode *Double Exponential Smoothing* (DES) menggunakan metrik evaluasi MAE, MSE, RMSE, dan R^2 untuk menilai tingkat akurasi prediksi.
3. Mengidentifikasi variabel yang paling berpengaruh terhadap fluktuasi harga jagung melalui analisis *feature importance*, guna mendukung

pengambilan keputusan yang lebih tepat dalam pengelolaan harga komoditas pertanian.

1.3.2. Manfaat Penelitian

1.3.2.1. Manfaat Teoritis

1. Memberikan kontribusi terhadap pengembangan literatur di bidang *machine learning* dan analisis data pertanian, khususnya penerapan algoritma *Random Forest Regression* pada kasus prediksi harga komoditas.
2. Memperkuat kajian empiris tentang efektivitas pendekatan ensemble learning dalam menangani data deret waktu non-linear di sektor pertanian.
3. Menjadi referensi bagi penelitian selanjutnya yang ingin mengintegrasikan metode *machine learning* dengan model statistik klasik untuk meningkatkan akurasi prediksi harga komoditas pangan..

1.3.2.2. Manfaat Praktis

1. Memberikan alat bantu prediksi harga yang akurat bagi petani, pelaku pasar, dan pemerintah daerah dalam merencanakan produksi, distribusi, serta strategi stabilisasi harga jagung.
2. Menjadi dasar pertimbangan bagi Dinas Pertanian dan instansi terkait dalam perumusan kebijakan intervensi pasar, subsidi, dan tata niaga jagung berbasis data.
3. Mendukung digitalisasi sektor pertanian melalui implementasi sistem prediksi harga berbasis *Streamlit* yang mudah diakses dan dapat dimanfaatkan oleh berbagai pihak.

1.4. Batasan Masalah

Agar penelitian lebih terarah dan fokus, batasan masalah yang ditetapkan adalah sebagai berikut:

1. Komoditas yang dianalisis terbatas pada harga jagung tingkat konsumen di pasar tradisional Kabupaten Sampang. Komoditas ini dipilih karena memiliki peran strategis dalam ketahanan pangan daerah dan kontribusi besar terhadap ekonomi pertanian lokal.

2. Data yang digunakan berupa data historis harga jagung periode Januari 2022 – Desember 2024 yang diperoleh dari SISKAPERBAKO Jawa Timur.
3. Lokasi penelitian difokuskan pada dua pasar utama, yaitu Pasar Rongtengah dan Pasar Srimangunan, yang dipilih karena mewakili pusat aktivitas perdagangan dan distribusi jagung di Kabupaten Sampang.
4. Algoritma utama yang digunakan adalah *Random Forest Regression (RFR)* dengan pembandingan *Double Exponential Smoothing (DES)* sebagai metode peramalan klasik untuk menilai tingkat keunggulan model berbasis *machine learning*.
5. Evaluasi kinerja model dilakukan menggunakan empat metrik standar regresi, yaitu *Mean Absolute Error (MAE)*, *Mean Squared Error (MSE)*, *Root Mean Squared Error (RMSE)*, dan *Coefficient of Determination (R²)*.
6. Implementasi sistem prediksi diwujudkan dalam bentuk prototype aplikasi berbasis Streamlit, yang menampilkan hasil prediksi harga, visualisasi tren, dan fitur input data sederhana untuk memudahkan pengguna dari kalangan petani, pedagang, maupun dinas pertanian daerah.

1.5. Sistematika Penulisan Skripsi

Sistematika penulisan skripsi ini disusun untuk memberikan gambaran menyeluruh mengenai alur penelitian, mulai dari perumusan masalah hingga kesimpulan. Adapun pembagian sistematika penulisan skripsi sebagai berikut:

BAB I PENDAHULUAN

Berisi latar belakang, perumusan masalah, tujuan dan manfaat, batasan masalah, serta sistematika penulisan skripsi.

BAB II KAJIAN PUSTAKA

Menjelaskan teori-teori yang mendasari penelitian, seperti konsep harga konsumen, *machine learning*, algoritma *Random Forest Regression*,

Double Exponential Smoothing, evaluasi model, *feature importance*, serta penelitian terdahulu yang relevan.

BAB III METODOLOGI PENELITIAN

Menguraikan jenis penelitian, sumber dan karakteristik data, tahapan penelitian berdasarkan metode CRISP-DM, proses pengolahan data, pemodelan, evaluasi, serta rancangan sistem prediksi berbasis *Streamlit*.

BAB IV HASIL DAN PEMBAHASAN

Menyajikan hasil analisis data, proses pemodelan, evaluasi performa model, interpretasi *feature importance*, perbandingan metode, serta pembahasan hasil prediksi terkait fenomena lapangan dan penelitian sebelumnya.

BAB V PENUTUP

Memuat kesimpulan penelitian dan saran untuk pengembangan penelitian atau penerapan sistem prediksi harga ke depan.