

BAB I

PENDAHULUAN

1.1 Latar Belakang

Indonesia merupakan negara agraris dengan subsektor perkebunan sebagai penopang penting perekonomian nasional. Komoditas strategis seperti kelapa sawit, karet, kopi, kakao, dan tebu menyumbang devisa ekspor nonmigas sekaligus menjadi sumber mata pencaharian jutaan penduduk [1]. Data dari Kementerian Pertanian menunjukkan bahwa luas areal perkebunan besar terus meningkat, dengan kontribusi signifikan terhadap Produk Domestik Bruto (PDB) sektor pertanian [2].

Namun, subsektor ini menghadapi tantangan besar akibat perubahan iklim global. Pergeseran pola curah hujan, peningkatan suhu ekstrim, serta kejadian cuaca ekstrim terbukti mempengaruhi produktivitas perkebunan [3]. Variabilitas curah hujan dan perubahan suhu memiliki pengaruh langsung terhadap ketersediaan air dan produktivitas lahan [4]. Fluktuasi curah hujan yang ekstrim dapat menyebabkan penurunan hasil pertanian karena terganggunya keseimbangan air tanah dan efisiensi penyerapan nutrisi tanaman [5]. Pada tebu, distribusi curah hujan yang tidak seimbang mengurangi kuantitas sekaligus kualitas rendemen gula [6], sedangkan pada kakao, variabilitas curah hujan berdampak negatif terhadap stabilitas produksi [7].

Upaya prediksi hasil produksi sebelumnya masih banyak menggunakan model regresi linier atau pendekatan statistik klasik [8]. Meskipun sederhana, metode tersebut terbatas dalam menangkap hubungan non-linear antara faktor iklim dan produksi. Seiring perkembangan teknologi, algoritma machine learning mulai diterapkan. Model seperti *Random Forest* dan *XGBoost* mampu memberikan hasil yang lebih akurat dalam memprediksi produktivitas pertanian [9].

Salah satu algoritma yang menonjol adalah Category Boosting (CatBoost), bagian dari keluarga Gradient Boosting Decision Trees (GBDT). Algoritma ini unggul dalam menangani data tabular dengan kombinasi fitur numerik dan kategorikal tanpa memerlukan proses manual encoding, serta terbukti stabil dan akurat [10]. Kemampuan ini sangat relevan dengan penelitian ini, karena model

harus memproses variabel kategorikal krusial seperti jenis tanaman (mencakup karet, minyak sawit, kopi, Gula Tebu, Teh) dan bulan, yang masing-masing memiliki pola produksi dan karakteristik musiman yang unik. Dengan menangani fitur-fitur ini secara *native*, Category Boosting dapat menghindari proses *one-hot encoding* manual yang berisiko memperbanyak dimensi data. Category Boosting juga telah terbukti di penelitian lain mampu memprediksi kebutuhan pupuk nitrogen dengan tingkat akurasi tinggi ($R^2 \approx 0,98$) dan efektif untuk estimasi biomassa hutan [11].

Sebagian besar penelitian sebelumnya masih fokus pada evaluasi model menggunakan pembagian data acak (random split) atau k-fold cross validation. Pendekatan tersebut seringkali memberikan hasil optimistis namun kurang relevan dalam konteks prediksi karena model dapat “melihat” pola dari masa depan yang seharusnya tidak tersedia saat prediksi dilakukan [12]. Evaluasi berbasis waktu seperti temporal split atau *Walk Forward* lebih realistis untuk menilai kemampuan model dalam memprediksi periode mendatang [13].

Dalam penelitian ini, Category Boosting diterapkan untuk memprediksi produksi perkebunan besar di Indonesia dengan mempertimbangkan dua variabel iklim utama, yaitu curah hujan dan suhu. Penelitian ini tidak hanya menguji performa model pada data historis, tetapi juga melakukan analisis sensitivitas untuk menguji ketangguhan model. Simulasi ini dilakukan dengan memproyeksikan produksi hingga tahun 2027 menggunakan pendekatan *Walk Forward forecasting* di bawah skenario uji stres (stress test) berupa fluktuasi iklim ekstrim $\pm 10\%$ dari rata-rata historis [14]. Selain itu, penelitian ini menerapkan optimasi hiperparameter menggunakan *Random Search* untuk meningkatkan akurasi model. Evaluasi model dilakukan melalui dua tahap eksperimen, yaitu pembagian berbasis waktu (temporal split), dan model ter optimasi (*Random Search optimized*).

Penerapan Category Boosting untuk prediksi produksi perkebunan besar di Indonesia masih jarang dilakukan. Dengan memanfaatkan data produksi perkebunan dari BPS dan data curah hujan dari BMKG, penelitian ini berupaya membangun model prediksi berbasis Category Boosting yang diharapkan mampu

mendukung perencanaan panen, manajemen rantai pasok, serta kebijakan adaptasi iklim berbasis data di sektor perkebunan Indonesia.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang diatas maka rumusan masalah yang didapat untuk penelitian ini adalah :

1. Bagaimana penerapan algoritma Category Boosting dalam memprediksi produksi perkebunan besar di Indonesia dengan memanfaatkan data historis produksi, curah hujan, suhu dan luas lahan ?
2. Bagaimana tingkat akurasi model Category Boosting Memprediksi Produksi Perkebunan Besar di Indonesia
3. Bagaimana Pengaruh Optimasi hyperparameter menggunakan *Random Search* terhadap performa model Category Boosting ?
4. Bagaimana hasil penerapan Category Boosting untuk prediksi produksi perkebunan hingga tahun 2027 ?

1.3 Batasan Masalah

Berdasarkan uraian rumusan masalah diatas maka batasan masalah yang didapat untuk penelitian ini supaya tidak keluar dari topik bahasan adalah :

1. Penelitian hanya difokuskan pada prediksi produksi perkebunan besar nasional (Karet kering, minyak sawit, kopi, teh, gula tebu) di Indonesia.
2. Data yang digunakan berasal dari Badan Pusat Statistik (BPS) untuk data produksi perkebunan, luas lahan dan BMKG untuk data curah hujan, dengan rentang tahun 2009–2024.
3. Variabel independen yang digunakan meliputi curah hujan rata-rata bulanan sebagai variabel iklim, serta tahun, bulan, luas lahan, jenis tanaman sebagai variabel temporal dan kategorikal.
4. Algoritma yang digunakan adalah Category Boosting sebagai model utama.
5. Evaluasi model dilakukan menggunakan metrik MAE, MSE, RMSE, dan R^2 .

6. Simulasi prediksi 2025-2027 dilakukan sebagai uji stres (stress test) model, menggunakan data iklim yang disimulasikan dengan skenario fluktuasi $\pm 10\%$ dari rata-rata historis, dan bukan merupakan prediksi klimatologi kompleks.
7. Prediksi 2025–2027 dilakukan menggunakan metode Walk Forward (prediksi bertahap) berdasarkan hasil model terbaik

1.4 Tujuan Penelitian

Berdasarkan penyampaian rumusan masalah dan batasan masalah diatas maka tujuan dari penulisan ini adalah :

1. Menerapkan algoritma Category Boosting untuk memprediksi produksi perkebunan besar di Indonesia dengan memanfaatkan data historis produksi, curah hujan, suhu, dan luas lahan.
2. Mengevaluasi tingkat akurasi dan performa model Category Boosting dalam memprediksi produksi perkebunan besar di Indonesia.
3. Mengoptimalkan parameter model Category Boosting menggunakan metode *Random Search* untuk meningkatkan akurasi prediksi.
4. Menerapkan model Category Boosting untuk melakukan prediksi produksi perkebunan besar di Indonesia hingga tahun 2027.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut:

1. Bagi Penulis
Penelitian ini memberikan pengalaman langsung dalam menerapkan algoritma Category Boosting pada data tabular di bidang pertanian. Penulis memperoleh pemahaman teknis dan praktis mengenai pemodelan prediktif berbasis machine learning, serta keterampilan dalam mengolah dan menganalisis data iklim dan pertanian secara terpadu.
2. Bagi Masyarakat dan Sektor Pertanian
Penelitian ini diharapkan dapat memberikan kontribusi nyata bagi masyarakat, khususnya para pelaku sektor pertanian yang sangat bergantung pada informasi produksi untuk merencanakan aktivitas

budidaya. Model prediksi yang dihasilkan dapat digunakan sebagai acuan dalam memperkirakan ketersediaan komoditas pada periode mendatang, sehingga petani dapat menentukan waktu tanam, penggunaan pupuk, serta strategi pengelolaan lahan dengan lebih tepat. Selain itu, hasil prediksi yang lebih akurat dapat membantu penyuluh pertanian dalam memberikan rekomendasi yang berbasis data kepada petani, bukan hanya berdasarkan pengalaman atau pola historis yang tidak selalu stabil akibat perubahan iklim. Bagi pelaku industri perkebunan dan rantai pasok, informasi ini dapat digunakan untuk mempersiapkan proses distribusi, pengolahan hasil, serta perencanaan kapasitas produksi agar lebih efisien.