

PREDIKSI OMZET PEMASARAN TANAMAN HIAS MENGGUNAKAN *GRADIENT BOOSTING TREE* DI DKI JAKARTA TAHUN 2021

Indri Nur Sukmawati¹, Universitas Koperasi Indonesia

Mulyati Eka Saputri², Universitas Koperasi Indonesia

Teni Deinarosa H³, Universitas Koperasi Indonesia

Ashafa Multazam⁴, Universitas Koperasi Indonesia

e-mail : sukmawatiindri6@gmail.com

Abstrak. Penelitian ini bertujuan untuk memprediksi omzet pemasaran tanaman hias di DKI Jakarta tahun 2021 menggunakan algoritma *Gradient Boosting Tree* (GBT). Dataset berupa data penjualan bulanan tanaman hias diproses melalui tahapan pembersihan data, transformasi logaritmik, *encoding* kategorikal, serta pembagian data latih dan uji. Tiga model regresi dibandingkan, yaitu *Random Forest Regression* (RFR), *Gradient Boosting Regression* (GBR), dan *Gradient Boosting Tree* (GBT), dengan evaluasi matriks menggunakan *Mean Squared Error* (MSE) dan *Koefisien Determinasi* (R^2). Hasil menunjukkan bahwa *Gradient Boosting Tree* (GBT) memberikan performa terbaik dengan MSE sebesar 0.1024 dan R^2 sebesar 0.9983. *Scatter plot* memperlihatkan distribusi prediksi *Gradient Boosting Tree* (GBT) yang paling mendekati nilai aktual. Simulasi prediksi bulan ke-7 menunjukkan bahwa Anggrek Bulan, *Phalaenopsis*, dan Melati merupakan komoditas dengan omzet tertinggi. Temuan ini menunjukkan bahwa pemodelan prediktif berbasis *Gradient Boosting Tree* (GBT) efektif untuk mendukung perencanaan pemasaran hortikultura secara lebih akurat, efisien, dan berbasis data.

Kata Kunci: *Gradient Boosting Tree, prediksi omzet, tanaman hias, machine learning, data mining*

Abstract. This study aims to predict the sales revenue (omzet) of ornamental plants in DKI Jakarta in 2021 using the *Gradient Boosting Tree* (GBT) algorithm. The dataset, comprising monthly sales data of various ornamental plant commodities, underwent preprocessing steps including missing value removal, logarithmic transformation of the target variable, categorical encoding, and a train-test data split. Three regression models: *Random Forest Regression* (RFR), *Gradient Boosting Regression* (GBR), and *Gradient Boosting Tree* (GBT) were evaluated using *Mean Squared Error* (MSE) and the *Coefficient of Determination* (R^2).

Results indicate that *Gradient Boosting Tree* (GBT) performed the best, with an MSE of 0.1024 and R^2 of 0.9983. Scatter plots showed that *Gradient Boosting Tree* (GBT) predictions were closely aligned with actual values. A simulation for month 7 revealed that Anggrek Bulan (*Phalaenopsis amabilis*), *Phalaenopsis*, and Melati (Jasmine) were among the highest revenue commodities. These findings demonstrate that predictive modeling using *Gradient Boosting Tree* (GBT) is an effective strategy to support more accurate, efficient, and data driven planning in ornamental plant marketing.

Keywords: *Gradient Boosting Tree, revenue prediction, ornamental plants, machine learning, data mining*

PENDAHULUAN

Salah-satu komoditas hortikultura yang menjadi primadona saat ini adalah tanaman hias, karena nampak keindahan fisiknya dan dapat menambah nilai estetika lingkungan.

Keanekaragaman jenis tanaman hias menjadikannya cocok digunakan sebagai elemen dekoratif, baik untuk keperluan interior maupun eksterior. Selain memiliki nilai estetis, sektor bisnis tanaman hias juga memberikan kontribusi terhadap Produk Domestik Bruto (PDB). Daerah Khusus Ibukota (DKI) Jakarta menjadi salah satu wilayah strategis dalam distribusi dan pemasaran tanaman hias, didukung oleh tingginya mobilitas masyarakat serta akses jaringan pemasaran yang luas. Di wilayah ini, terdapat berbagai pelaku usaha tanaman hias, baik berskala kecil, menengah, maupun besar, yang tersebar di beberapa sentra utama di wilayah Jakarta Selatan, Barat, Pusat, Utara, dan Timur.

Pada tahun 2021, tren pemasaran tanaman hias di DKI Jakarta mengalami dinamika yang cukup signifikan seiring perubahan pola konsumsi masyarakat akibat Covid-19. Fenomena aktivitas berkebun di rumah yang meningkat selama masa pembatasan sosial telah mendorong lonjakan permintaan, sementara pelonggaran PPKM di pertengahan tahun menyebabkan pola pemasaran kembali berubah. Kondisi ini menyebabkan omzet penjualan cenderung fluktuatif, sehingga diperlukan prediksi omzet bulanan untuk membantu pelaku usaha dalam merencanakan strategi bisnis yang adaptif.

Penerapan teknologi *data mining* dan *machine learning* menjadi alternatif solusi yang dapat digunakan untuk membangun sistem prediksi omzet tanaman hias berdasarkan data historis, seperti jenis tanaman, volume penjualan, dan lokasi pemasaran. Penelitian terdahulu menunjukkan bahwa algoritma *Random Forest* memiliki kemampuan prediksi yang tinggi dalam domain pertanian dan penjualan produk hortikultura, penelitian menunjukkan bahwa *Random Forest* mampu menghasilkan prediksi hasil pertanian dengan akurasi tinggi (Padma & Sinha, 2023). Selain *Random forest* kita memilih algoritma *Gradient Boosting Regression* sebagai metode prediksi karena kemampuannya dalam menangani data *non-linier*, mendeteksi pola tersembunyi dalam data historis, serta menghasilkan prediksi yang lebih presisi (Khan, Mishra, & Baranidharan, 2020). Sementara itu, (Rizkallah, 2025) menyatakan bahwa *Gradient Boosting Tree* (GBT) merupakan salah satu algoritma paling efektif untuk memprediksi data regresi dengan karakteristik variabel yang kompleks dan *non-linier*. Dalam penelitiannya, *Gradient Boosting Tree* (GBT) terbukti mampu memberikan performa prediksi yang unggul dibandingkan model tunggal lainnya, dengan rata-rata peningkatan akurasi hingga 75% pada berbagai dataset regresi.

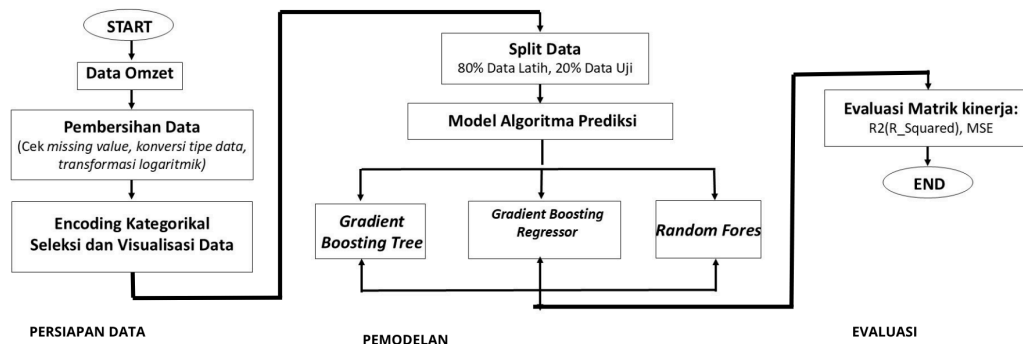
Dalam penelitian ini dilakukan prediksi omzet bulanan menggunakan pendekatan *machine learning* berbasis data historis penjualan komoditas. Data yang digunakan merupakan data *time series* penjualan komoditas selama periode tertentu yang telah melalui tahap pra proses, seperti cek *missing value*, *encoding kategorikal*, seleksi dan visualisasi data. Penelitian ini membandingkan kinerja beberapa algoritma regresi, yaitu *Random Forest Regression* (RFR), *Gradient Boosting Regression* (GBR) dan *Gradient Boosting Tree* (GBT). Pemilihan algoritma *Gradient Boosting Tree* (GBT) didasarkan pada kemampuannya dalam menangani data *non-linear* serta kompleksitas variabel yang mempengaruhi penjualan. Evaluasi kinerja model dilakukan menggunakan matriks evaluasi *Mean Squared Error* (MSE) dan Koefisien Determinasi (R^2)

Berdasarkan latar belakang permasalahan diatas, rumusan masalah dalam penelitian ini adalah bagaimana proses persiapan dan pengolahan data penjualan komoditas untuk peramalan omset, bagaimana perancangan model prediksi yang sesuai, serta bagaimana mengukur kinerja model yang digunakan dalam menghasilkan prediksi omzet bulanan.

Penelitian ini bertujuan untuk mengidentifikasi metode regresi yang paling sesuai dalam melakukan prediksi omzet berdasarkan data deret waktu penjualan, dengan fokus pada pencapaian tingkat akurasi yang optimal.

METODE PENELITIAN

Pada bab ini akan menjelaskan metode proses tahapan untuk membuat model peramalan prediksi omzet pemasaran tanaman hias dan bunga di DKI Jakarta. Tahapan ini terdiri dari 3 sub bab yaitu persiapan data, pemodelan, dan evaluasi.



Gambar 1. Alur Penelitian

1) Persiapan data

Tahap awal dari persiapan data adalah melakukan pengumpulan data. Dataset penelitian ini menggunakan dataset jumlah omzet pemasaran tanaman hias di provinsi DKI Jakarta Tahun 2021. Setelah mendapatkan data, langkah pertama dalam pengolahan data adalah memeriksa keberadaan nilai kosong (*missing value*), khususnya pada kolom omzet yang menjadi target prediksi, seluruh baris dengan nilai kosong pada kolom omzet dihapus untuk menjaga kualitas dan integrasi data.

Selanjutnya, kolom omzet dikonversi ke tipe data numerik (*float*) agar dapat dilakukan pengolahan matematis. konversi ini penting karena algoritma *machine learning* hanya dapat memproses data dalam bentuk numerik, sementara data mentah sering kali tersimpan dalam format string atau mengandung simbol tertentu. Untuk mengurangi *skewness* dan pengaruh *outlier* pada data omzet, transformasi logaritmik menggunakan fungsi $\log_{10}()$ diterapkan untuk menstabilkan variansi dan memperkecil dampak outlier terhadap model. Hasil dari transformasi ini disimpan dalam kolom baru bernama `log_omzet` yang digunakan sebagai target dalam proses pelatihan model.

Beberapa kolom kategorikal seperti `lokasi_omzet_penjualan`, `jenis_komoditi`, dan `satuan` juga diolah menggunakan teknik label encoding agar dapat diproses oleh algoritma pemodelan yang hanya menerima input numerik, kecuali untuk model *Gradient Boosting Regression* yang mampu mengolah data kategorikal secara langsung. Data juga divisualisasikan menggunakan *histogram*, *boxplot*, dan *scatter plot* untuk mengeksplorasi distribusi, mendeteksi *outlier*, dan memahami pola hubungan antar variabel. Fitur-fitur yang tidak relevan terhadap proses pemodelan,

seperti *periode_data* dan omzet asli (setelah dikonversi ke log) , dihapus dari input variabel X agar tidak terjadi kebocoran data (*data leakage*).

2) Pemodelan

Setelah data selesai diproses, langkah selanjutnya adalah pemodelan. Dataset dibagi menjadi dua bagian, yaitu data latih sebesar 80% dan data uji sebesar 20% menggunakan fungsi *train_test_split* dari library *scikit-learn*. Tahapan selanjutnya adalah pembanguna model prediksi menggunakan tiga algoritma regresi, yaitu *Random Forest Regression* (RFR), *Gradient Boosting Tree* (GBT), dan *Gradient Boosting Regression* (GBR).

Model *Random Forest Regression* dibangun menggunakan parameter standar dari pustaka *scikit-learn*, sedangkan model *Gradient Boosting Tree* menggunakan beberapa parameter, yakni: *n_estimators* = 500, *learning_rate* = 0.05, *max_depth* = 5, dan *subsample* = 0.8. Adapun model *Gradient Boosting Regression* diinisialisasi dengan parameter *iterations* = 1000, *learning_rate* = 0.05, *depth* = 6.

Setiap model dilatih menggunakan data latih (80% dari dataset), kemudian digunakan untuk melakukan prediksi nilai *log_omzet* pada data uji (20%). Hasil prediksi selanjutnya dikonversi kembali ke skala asli omzet dengan menggunakan fungsi *expm1()* agar dapat dilakukan interpretasi dan analisis lebih lanjut.

Sebagai tambahan, dilakukan simulasi prediksi omzet pada bulan ke-7 menggunakan data *dummy* yang dibuat berdasarkan informasi penjualan pada bulan ke-6. Data ini diolah dengan mengganti nilai bulan dan mengisi ulang volume berdasarkan rata-rata per jenis komoditas, kemudian digunakan sebagai input untuk memprediksi omzet ke depan.

3) Evaluasi Model

Untuk mengukur kinerja model, dua matrik evaluasi utama digunakan yaitu *R-squared* (R^2) dan *Mean Square Error* (MSE). MSE digunakan untuk mengukur besarnya kesalahan rata-rata kuadrat antara nilai prediksi dan aktual, di mana nilai yang lebih kecil menunjukkan prediksi yang lebih akurat. Sementara itu, nilai R^2 menunjukkan proporsi variansi dalam variabel target yang dapat dijelaskan oleh model, semakin mendekati 1, semakin baik model menjelaskan data.

Selain itu, visualisasi hasil prediksi dilakukan dengan menampilkan *scatter plot* antara nilai aktual dan prediksi, serta histogram distribusi error untuk menilai penyebaran kesalahan model. Visualisasi tambahan juga ditampilkan untuk memperlihatkan 10 komoditas tanaman hias dengan prediksi omzet tertinggi pada bulan ke-7.

HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil dari penerapan model prediksi terhadap data omzet pemasaran tanaman hias di Provinsi DKI Jakarta tahun 2021. Proses analisis mencakup tahapan evaluasi performa model regresi, analisis visual terhadap distribusi hasil prediksi, serta identifikasi komoditas dengan potensi tertinggi berdasarkan hasil pemodelan. Evaluasi

dilakukan terhadap tiga algoritma regresi yang digunakan, yaitu *Random Forest Regression*, *Gradient Boosting Tree* (GBT), dan *Gradient Boosting Regression* (GBR).

Hasil

1. Deskripsi dan Persiapan Data

Data penelitian ini mencakup informasi omzet penjualan tanaman hias di wilayah DKI Jakarta selama tahun 2021, dengan variabel-variabel yang terdiri dari jenis komoditas, lokasi penjualan, satuan transaksi, serta nilai omzet bulanan. Tahapan awal dalam pengolahan data dilakukan dengan menghapus data yang tidak lengkap (*missing value*) pada variabel target, yaitu omzet, untuk memastikan keandalan dalam proses pemodelan.

Nilai pada variabel target omzet selanjutnya ditransformasi ke dalam bentuk logaritmik menggunakan fungsi $\log_{1p}()$, guna mengatasi masalah distribusi data yang tidak normal serta mereduksi pengaruh nilai ekstrim (*outlier*). Beberapa atribut kategorikal, seperti jenis komoditi, lokasi penjualan, dan satuan, diubah menjadi bentuk numerik melalui teknik *label encoding*. Data yang telah diproses selanjutnya dibagi ke dalam dua bagian, yakni data latih sebesar 80% dan data uji sebesar 20%, untuk keperluan pelatihan dan evaluasi model prediksi.

2. Evaluasi Kinerja Model

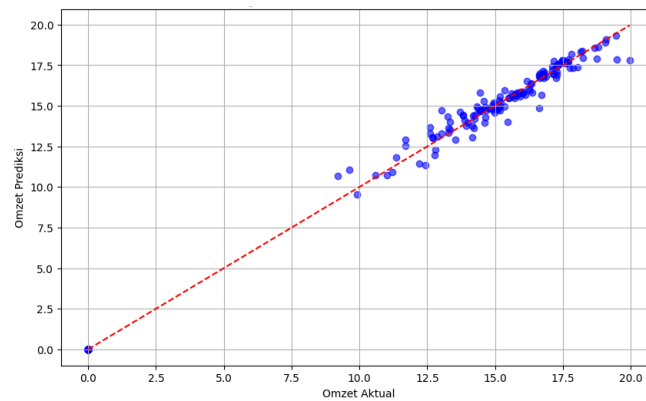
Model-model regresi dilatih dan dievaluasi menggunakan metrik *Mean Squared Error* (MSE) dan *Koefisien Determinasi* (R^2). Hasil evaluasi ditunjukkan pada Tabel 1 berikut.

Tabel 1. Hasil Evaluasi Model Regresi

METODE	MSE	R^2
<i>Random Forest</i>	0.1763	0.9971
<i>Gradient Boosting Tree</i>	0.1024	0.9983
<i>Gradient Boosting Regression</i>	0.2039	0.9967

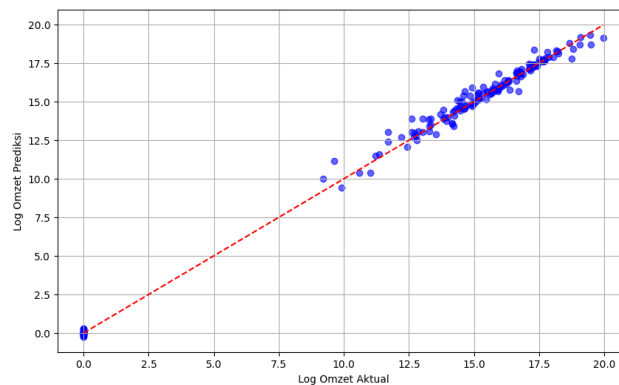
Hasil evaluasi menunjukkan bahwa model GBT memiliki nilai MSE paling rendah dan R^2 paling tinggi, menandakan performa prediktif terbaik. Model *Random Forest Regression* (RFR) menduduki posisi kedua, sedangkan *Random Forest Regression* (GBR) menunjukkan hasil yang cukup baik namun kurang akurat dibanding dua model lainnya.

3. Visualisasi Hasil Prediksi



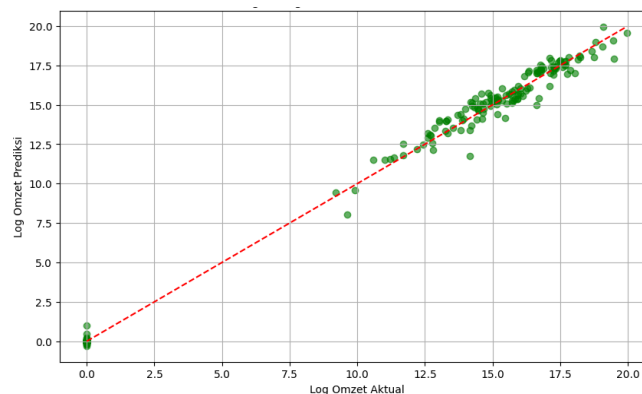
Gambar 2. Perbandingan Omzet Aktual vs Prediksi – *Random Forest*

Gambar ini menunjukkan hubungan antara nilai aktual dan nilai prediksi dari model *Random Forest Regression* (RFR). Titik-titik biru mewakili prediksi pada masing-masing data, sedangkan garis merah putus-putus adalah garis ideal (prediksi = aktual). Distribusi titik yang mendekati garis diagonal menunjukkan bahwa model memiliki prediksi yang cukup akurat. Meskipun demikian, terdapat sedikit deviasi pada prediksi dengan nilai omzet tinggi.



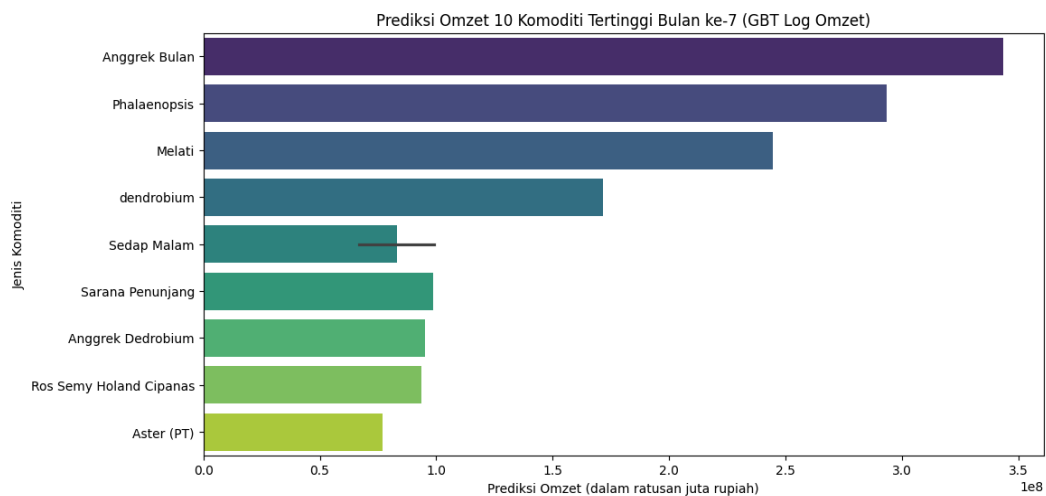
Gambar 3. Perbandingan Log Omzet Aktual vs Prediksi – *Gradient Boosting Tree*

Scatter plot menunjukkan bahwa prediksi dari *Gradient Boosting Tree* hampir seluruhnya berada di sekitar garis ideal. Hal ini mengindikasikan bahwa model berhasil mempelajari pola data dengan baik, terutama pada nilai omzet tinggi maupun rendah. Ini sejalan dengan nilai MSE terkecil dan R^2 tertinggi dari model ini.



Gambar 4. Perbandingan Log Omzet Aktual vs Prediksi – *Gradient Boosting Regression*

Gambar ini menunjukkan bahwa performa *Gradient Boosting Regression* cukup baik namun tidak seakurat *Gradient Boosting Tree*. Beberapa titik prediksi menyimpang dari garis ideal, terutama pada rentang omzet menengah. Meski demikian, secara keseluruhan, prediksi tetap berada dalam kisaran yang dapat diterima.



Gambar 5. Prediksi Omzet 10 Komoditas Tertinggi Bulan ke-7 (GBT)

Visualisasi bar chart menampilkan sepuluh komoditas dengan prediksi omzet tertinggi pada bulan ke-7. Komoditas Anggrek Bulan, *Phalaenopsis*, dan Melati menjadi tiga jenis dengan prediksi omzet terbesar. Temuan ini dapat menjadi acuan strategis dalam alokasi sumber daya dan perencanaan distribusi komoditas hortikultura.

Pembahasan

1. Kinerja Model dan Interpretasi Statistik

Evaluasi kinerja model dilakukan menggunakan dua metrik utama, yakni *Mean Squared Error* (MSE) dan *Koefisien Determinasi* (R^2). Hasil menunjukkan bahwa *Gradient Boosting Tree* (GBT) memiliki performa terbaik dengan MSE sebesar 0.1024 dan R^2 sebesar 0.9983, mengungguli *Random Forest Regression* (MSE = 0.1763; R^2 = 0.9971) dan *Gradient Boosting Regression* (MSE = 0.2039; R^2 = 0.9967).

Secara visual, *scatter plot* antara nilai aktual dan prediksi menunjukkan bahwa *Gradient Boosting Tree* (GBT) menghasilkan distribusi titik yang paling mendekati garis ideal, menandakan akurasi tinggi dan stabilitas prediksi pada berbagai rentang omzet. *Bar chart* komoditas menunjukkan bahwa Anggrek Bulan, *Phalaenopsis*, dan Melati merupakan jenis dengan prediksi omzet tertinggi pada bulan ke-7.

Hasil ini sejalan dengan penelitian sebelumnya (Khan, Mishra, & Baranidharan, 2020) yang menunjukkan keunggulan *Gradient Boosting Tree* (GBT) dalam memodelkan data non-linear dan kompleks. Kombinasi nilai statistik dan visualisasi menguatkan bahwa *Gradient Boosting Tree* (GBT) merupakan model paling efisien dan akurat untuk prediksi omzet tanaman hias di DKI Jakarta.

2. Kesesuaian Temuan dengan Literatur

Hasil penelitian menunjukkan bahwa algoritma *Gradient Boosting Tree* (GBT) memberikan kinerja terbaik dalam memprediksi omzet pemasaran tanaman hias, dengan nilai MSE sebesar 0.1024 dan koefisien determinasi (R^2) sebesar 0.9983. Hal ini menandakan bahwa *Gradient Boosting Tree* (GBT) mampu menjelaskan lebih dari 99% variasi data omzet yang ada dalam dataset, serta menghasilkan galat prediksi yang sangat kecil. Temuan ini sejalan dengan penelitian yang dilakukan oleh (Rizkallah, 2025) yang menyatakan bahwa algoritma *Gradient Boosting* memiliki performa unggul dalam konteks prediksi hasil pertanian dan hortikultura. GBT terbukti efektif dalam menangani data yang memiliki hubungan *non-linier* dan kompleks antar variabel, seperti yang juga ditemukan dalam data penjualan komoditas tanaman hias ini.

Sebagai perbandingan, algoritma *Random Forest* juga menunjukkan performa prediktif yang sangat baik, dengan MSE sebesar 0.1763 dan R^2 sebesar 0.9971, namun masih berada di bawah *Gradient Boosting Tree* (GBT). Hasil ini konsisten dengan studi oleh (Padma et al., 2021), yang mengemukakan bahwa meskipun *Random Forest* memiliki stabilitas tinggi, algoritma boosting seperti *Gradient Boosting Tree* (GBT) cenderung memberikan hasil yang lebih presisi dalam konteks regresi yang kompleks. Sementara itu, *Gradient Boosting Regression* (GBR) yang pada penelitian ini disebut sebagai model ketiga, mencatat nilai MSE tertinggi yaitu 0.2039, meskipun R^2 masih sangat tinggi yakni 0.9967. Hal ini menunjukkan bahwa meskipun secara umum ketiga model mampu memodelkan data dengan baik, terdapat perbedaan dalam ketelitian model saat melakukan prediksi pada data uji.

Secara keseluruhan, temuan penelitian ini memperkuat literatur yang telah ada, bahwa algoritma *ensemble* berbasis *boosting* sangat direkomendasikan untuk prediksi data penjualan komoditas yang bersifat musiman, fluktuatif, dan multivariat, seperti yang umum ditemukan pada sektor agribisnis hortikultura.

3. Pengaruh Parameter terhadap Akurasi Model

Kinerja tinggi dari model *Gradient Boosting Tree* (GBT) dipengaruhi oleh kombinasi parameter yang optimal, yakni $n_estimators = 500$, $learning_rate = 0.05$, $max_depth = 5$, dan $subsample = 0.8$. Penggunaan $learning_rate$ rendah memungkinkan pembelajaran bertahap dan menghindari *overfitting*, sementara kedalaman pohon yang moderat menjaga keseimbangan antara kompleksitas dan generalisasi model.

Subsampling turut mengurangi korelasi antar pohon dan meningkatkan ketahanan model terhadap variasi data. Temuan ini sejalan dengan (Budholiya, Shrivastava & Sharma, 2020) yang menekankan pentingnya tuning parameter dalam meningkatkan akurasi model boosting, khususnya pada data prediksi agrikultur yang bersifat kompleks dan *non-linear*.

4. Implikasi Hasil terhadap Strategi Pemasaran

Hasil prediksi yang akurat dari model *Gradient Boosting Tree* memiliki implikasi praktis yang signifikan terhadap strategi pemasaran komoditas tanaman hias, khususnya di wilayah DKI Jakarta yang memiliki dinamika permintaan tinggi. Visualisasi prediksi omzet bulan ke-7 menunjukkan bahwa komoditas seperti Anggrek Bulan, *Phalaenopsis*, dan Melati memiliki potensi nilai ekonomi tertinggi.

Informasi ini dapat dimanfaatkan oleh pelaku usaha dan pemangku kebijakan dalam menyusun perencanaan produksi, distribusi, dan promosi yang lebih terarah. Prediksi omzet yang presisi memungkinkan pengalokasian sumber daya secara efisien, pengendalian stok yang tepat waktu, serta penentuan harga yang kompetitif.

Selain itu, penerapan model prediktif berbasis data historis memungkinkan pelaku pasar untuk melakukan antisipasi terhadap fluktuasi musiman, mengoptimalkan strategi pemasaran berbasis lokasi dan jenis komoditas, serta merumuskan kebijakan jangka pendek dan menengah yang lebih responsif terhadap kebutuhan pasar.

SIMPULAN

Penelitian ini menunjukkan bahwa algoritma *Gradient Boosting Tree* (GBT) memberikan performa terbaik dalam memprediksi omzet pemasaran tanaman hias di DKI Jakarta tahun 2021, dengan nilai MSE sebesar 0.1024 dan R^2 sebesar 0.9983. Model ini unggul dibandingkan *Random Forest* dan *Gradient Boosting Regression* dalam hal akurasi dan stabilitas prediksi.

Hasil prediksi yang akurat, khususnya pada simulasi bulan ke-7, menunjukkan bahwa komoditas seperti Anggrek Bulan, *Phalaenopsis*, dan Melati memiliki potensi omzet tertinggi. Temuan ini dapat dimanfaatkan sebagai dasar dalam strategi pemasaran, perencanaan distribusi, serta pengambilan keputusan berbasis data di sektor agribisnis hortikultura.

Model *Gradient Boosting Tree* (GBT) terbukti efektif dalam mengelola data penjualan yang bersifat *non-linear*, fluktuatif, dan kompleks, serta direkomendasikan untuk diterapkan lebih lanjut dalam sistem peramalan komoditas sejenis.

DAFTAR PUSTAKA

Ardianto, A., Raharjo, A. B., & Purwitasari, D. (2022). *Random forest regression untuk prediksi produksi daya pembangkit listrik tenaga surya*. PLN Unit Pelaksana Pengatur Distribusi Makassar & Institut Teknologi Sepuluh Nopember. From <https://jurnal.unublitar.ac.id/index.php/briliant/article/view/1036>

Budholiya, K., Shrivastava, R., & Sharma, V. (2020). *Performance analysis of decision tree and ensemble methods for crop yield prediction*. Proceedings of the 2020 International Conference on Intelligent Engineering and Management (ICIEM), 91–95. doi: 10.1109/ICIEM48762.2020.9160130

Khan, R., Mishra, P., & Baranidharan, B. (2020). *Crop yield prediction using gradient boosting regression*. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 9(3), 123–127. From https://www.researchgate.net/publication/364050396_Crop_Yield_Prediction_using_Gradient_Boosting_Regression

Martoyo, A., Hendrawan, S. A., & Hesnanda, R. (2022). *Perencanaan strategi pemasaran tanaman hias Taman Anggrek Ragunan di masa pandemi Covid-19*. Jurnal

Sekretari dan Manajemen. From

<https://web.archive.org/web/20220907001346/https://ejournal.bsi.ac.id/ejurnal/index.php/widyacipta/article/download/12379/pdf>

Padma, T., & Sinha, D. (2023). *Crop yield prediction using improved random forest*. ITM Web of Conferences, 56, 02007. From

https://www.itm-conferences.org/articles/itmconf/pdf/2023/06/itmconf_icdsac2023_02007.pdf

Rizkallah, L. W. (2025). *Enhancing the performance of gradient boosting trees on regression problems*. Journal of Big Data, 12(35). doi: 10.1186/s40537-025-01071-3