Jurnal Strategi dan Inovasi Bisnis (JSIB)
Volume 2 Nomor 1 Tahun 2026
E-ISSN: 3109-7715

Implementasi Analitika Prediktif untuk Peramalan Permintaan Produk Cat TEFA Menggunakan Random Forest dan Regresi Linier Berganda

Dwiga Agus Nugroho

Sekolah Menengah Kejuruan SMTI Bandar Lampung, Indonesia

ABSTRAK

Latar Belakang: Unit bisnis *Teaching Factory* di SMK SMTI Bandar Lampung menghadapi tantangan operasional dalam menyelaraskan volume permintaan pasar dengan kapasitas produksi cat yang terbatas pada 5.000 Kg per bulan. Selama ini, penentuan kuota produksi didominasi oleh insting manajemen, sehingga meningkatkan risiko inefisiensi seperti penumpukan stok atau kehabisan bahan baku.

Tujuan: Untuk mengimplementasikan dan membandingkan model analitika prediktif menggunakan perangkat lunak *Altair AI Studio* (*RapidMiner*) untuk meramalkan volume penjualan produk cat secara akurat sehingga keputusan manajerial dapat berbasis data.

Metode: Penelitian kuantitatif ini menganalisis data operasional 18 bulan untuk empat varian produk cat. Pengaruh Harga, Kualitas, E-CRM, dan Citra Institusi dievaluasi terhadap Volume Penjualan. Analisis dilakukan dengan membandingkan algoritma *Multiple Linear Regression* (MLR) dan *Random Forest*.

Hasil: Pengujian *Linear Weights* menunjukkan variabel Citra Institusi tidak berpengaruh signifikan (p-value > 0.05). Sementara itu, variabel E-CRM terbukti memiliki pengaruh yang signifikan. Model *Random Forest* terbukti unggul dengan tingkat kesalahan terendah sebesar 435,52 Kg. Model ini mampu mengatasi multikolinearitas yang menyebabkan performa MLR menurun (RMSE 473,00 Kg).

Kesimpulan: Konsumen TEFA SMTI lebih mengutamakan faktor harga dan kualitas daripada citra institusi. Algoritma *Random Forest* sangat direkomendasikan untuk menjadi mesin peramal utama. Manajemen disarankan segera beralih ke penentuan kuota berbasis *Machine Learning*.

Kata Kunci: Peramalan*; Machine Learning; Random Forest; Teaching Factory;* Keputusan Berbasis Data

✉ Corresponding author: Dwiga Agus Nugroho
Email Address. dwiganugroho38@guru.smk.belajar.id
Doi: https://doi.org/10.69666/4aehk582

PENDAHULUAN

Unit bisnis Teaching Factory (TEFA) merupakan perwujudan nyata dari transformasi pendidikan vokasi di Indonesia yang berorientasi pada kemandirian institusi dan hilirisasi produk standar industri kepada masyarakat luas (Direktorat Sekolah Menengah Kejuruan, 2023). Sebagai sebuah unit produksi berskala Usaha Mikro Kecil dan Menengah (UMKM), TEFA SMK SMTI Bandar Lampung memiliki komoditas unggulan berupa produk cat otomotif dan dekoratif. Namun, operasional bisnis ini berhadapan dengan tantangan manajerial yang krusial, yaitu menyelaraskan fluktuasi permintaan pasar yang dinamis dengan kapasitas produksi pabrik yang memiliki batas maksimal sebesar 5.000 Kg per bulannya (Adnyana et al., 2025).

Berdasarkan observasi pada sistem operasional saat ini, praktik penentuan kuota produksi maklon cat di TEFA SMTI masih sangat bergantung pada pendekatan konvensional, yakni insting dan pengalaman empiris (*rule of thumb*) dari jajaran manajemen. Ketiadaan instrumen proyeksi berbasis data matematis ini meningkatkan risiko inefisiensi yang signifikan pada rantai pasok. Risiko tersebut termanifestasi dalam dua bentuk utama: terjadinya penumpukan barang jadi di gudang (*overstock*) ketika produksi berlebih, atau ketidakmampuan melayani permintaan konsumen (*stockout*) yang berujung pada hilangnya potensi pendapatan (*opportunity loss*) (Feizabadi, 2022) dan penurunan kepuasan pelanggan (Adnyana et al., 2025).

Dalam paradigma industri modern 4.0, keberadaan sistem peramalan (*forecasting*) yang akurat tidak lagi menjadi opsi, melainkan kebutuhan esensial untuk menjamin keberlangsungan hidup organisasi (Montgomery et al., 2024). Penelitian-penelitian terdahulu di bidang manajemen rantai pasok menegaskan bahwa integrasi data historis dengan sistem komputasi cerdas dapat mereduksi biaya operasional secara radikal (Witten et al., 2017). Oleh karena itu, penelitian ini dirancang untuk mengembangkan dan mengimplementasikan analitika prediktif menggunakan pendekatan Machine Learning guna memproyeksikan volume penjualan produk cat TEFA SMTI ke depan. Transisi dari "insting kepala pabrik" menuju "keputusan berbasis data" adalah standar wajib industri (Aamer et al., 2021).

Penelitian ini unik karena tidak hanya berfokus pada analisis tren tunggal, melainkan mengukur pengaruh empat variabel manajerial utama, yaitu: Kebijakan Harga, Kualitas Produk, Tata Kelola Electronic Customer Relationship Management (E-CRM), dan kekuatan Citra Institusi SMTI terhadap Volume Penjualan. Lebih jauh lagi, penelitian ini bertujuan untuk menguji komparasi efektivitas antara algoritma konvensional Multiple Linear Regression dengan algoritma tingkat lanjut Random Forest di dalam ekosistem perangkat lunak Altair AI Studio (RapidMiner). Hasil dari penelitian ini diharapkan dapat memberikan kerangka kerja baru bagi TEFA SMTI untuk melakukan transisi menyeluruh menuju sistem pengambilan keputusan berbasis data (data-driven decision making).

METODE PENELITIAN

Penelitian ini dilaksanakan dengan pendekatan eksperimental kuantitatif menggunakan partisipan berupa kumpulan rekaman data historis (dataset) penjualan dan dinamika variabel operasional dari unit bisnis TEFA SMK SMTI Bandar Lampung (Bougie & Sekaran, 2019; Sugiyono, 2022). Kriteria inklusi data yang digunakan mencakup rekam jejak transaksi selama 18 bulan terakhir (satu setengah tahun masa pembukuan) yang mencakup empat varian utama produk cat yang diproduksi oleh TEFA. Seluruh data ini diekstraksi secara internal dengan izin kepala pabrik dan telah melalui proses prapemrosesan data (*data preprocessing*) untuk membersihkan *missing value*s sebelum dimasukkan ke dalam lingkungan kerja perangkat lunak Altair AI Studio (RapidMiner) (Gulo & Machfud, 2026).
Desain arsitektur variabel dalam penelitian ini melibatkan satu Variabel Terikat (Y), yaitu Volume

Penjualan (dalam satuan Kilogram), yang diprediksi oleh empat Variabel Bebas (X). Keempat variabel bebas tersebut meliputi Harga (X1), Kualitas (X2), implementasi E-CRM (X3), dan Citra Institusi (X4). Instrumen pengukuran data sekunder ini dikompilasi langsung dari metrik operasional internal pabrik, survei kepuasan pelanggan B2B skala Likert (1-5), serta rekaman harga jual aktual di pasaran. Mengingat riset ini berfokus pada analitika Machine Learning, indikator reliabilitas model tidak diukur menggunakan koefisien Alpha Cronbach konvensional, melainkan menggunakan metrik validasi silang (Cross Validation) secara iteratif pada perangkat lunak, guna memastikan tidak terjadinya overfitting pada data latih atau *training data* (Witten et al., 2017).

Teknik analisis data dijalankan melalui dua tahapan utama menggunakan dua algoritma yang berbeda secara karakteristik.

- Pertama, algoritma Multiple Linear Regression (MLR) digunakan untuk tujuan inferensial diagnostik guna membaca nilai Linear Weights (setara p-value), sehingga peneliti dapat mengidentifikasi variabel mana yang gagal membuktikan signifikansinya terhadap penjualan (variabel jebakan).
- Kedua, algoritma Random Forest Regression difungsikan sebagai metode pembanding prediktif karena algoritma *ensemble* ini berbasis struktur pohon keputusan (Decision Tree) majemuk, yang secara teoritis sangat tangguh dalam menetralkan anomali multikolinearitas (korelasi berlebih antara variabel X1 dan X2).

Evaluasi performa akhir dari kedua model ini dibandingkan secara ketat berdasarkan besaran nilai Root Mean Square Error (RMSE) (Montgomery et al., 2024). Riset ini telah memenuhi kode etik penelitian dengan menjaga kerahasiaan data rasio keuangan institusi dan berfokus pada optimalisasi kapasitas pabrik.

HASIL DAN PEMBAHASAN

HASIL

Tahap pertama pengolahan data menggunakan algoritma Multiple Linear Regression di RapidMiner menghasilkan pembobotan fitur (Feature Weights) yang secara eksplisit memetakan variabel mana yang memiliki dorongan terkuat terhadap penjualan. Ditemukan bahwa variabel Citra Institusi SMTI (X4) memunculkan rasio probabilitas (p-value) yang melebihi ambang batas toleransi kesalahan (> 0.05). Hasil komputasi ini menjadikan variabel X4 tereliminasi secara statistik dari persamaan prediksi linier murni. Di sisi lain, variabel E-CRM (X3) menunjukkan nilai probabilitas (p-value) yang memenuhi ambang batas signifikansi (< 0.05) dengan bobot fitur yang positif. Hal ini mengindikasikan bahwa tata kelola hubungan pelanggan berbasis elektronik berkontribusi secara statistik terhadap retensi dan peningkatan volume pesanan dari konsumen.

Selanjutnya, pengujian bergeser pada pencarian model peramalan (*forecasting*) dengan error terkecil. Penggunaan regresi linier standar mengalami hambatan teknis akibat terdeteksinya gejala multikolinearitas pada dataset yakni kondisi di mana variabel Harga (X1) dan Kualitas (X2) saling berkorelasi terlalu kuat sehingga mengacaukan ekuilibrium regresi. Hal ini terbukti dari nilai Root Mean Square Error (RMSE) regresi linier yang membengkak di angka 473,00 Kg. Algoritma Random Forest kemudian diimplementasikan untuk menangani defisiensi tersebut. Ringkasan perbandingan komprehensif atas performa kedua model utama disajikan pada Tabel 1 di bawah ini.

Tabel 1. Komparasi Performa Evaluasi Model Machine Learning

Matriks Operasionalisasi Variabel		
Model Algoritma (RapidMiner)	**Metrik Kesalahan (RMSE)**	**Status Diagnostik dan Keterangan Model**
Multiple Linear Regression (MLR)	473,00 Kg	Mengalami bias tinggi akibat multikolinearitas (X1 & X2).
Random Forest Regression	435,52 Kg	Optimal. Meminimalisir deviasi error, tahan multikolinearitas.

Sumber: Data diolah menggunakan Altair AI Studio (2024)

PEMBAHASAN

Berdasarkan paparan hasil pengolahan data di atas, terdapat dua temuan krusial yang saling terkait dalam menjawab masalah manajerial di pabrik cat TEFA SMTI Bandar Lampung. Temuan pertama berkaitan dengan anatomi perilaku konsumen. Tereliminasinya variabel Citra Institusi SMTI (X4) membuktikan bahwa konsumen di sektor industri kimia (khususnya cat maklon) bergerak berdasarkan asas rasionalitas ekonomi murni. Entitas pembeli berstatus Business-to-Business (B2B) maupun Business-to-Consumer (B2C) sama sekali tidak menjadikan "nama besar sebuah sekolah vokasi" sebagai alasan utama pembelian. Konsumen mutlak memfokuskan keputusannya pada rasio Cost-to-Benefit, yakni keterjangkauan Harga (X1) yang harus selaras dengan ketangguhan Kualitas produk (X2). Temuan ini sejalan dengan teori perilaku konsumen rasional yang menegaskan bahwa pada produk fungsional komoditas, keunggulan teknis lebih diutamakan daripada sekadar merek atau afiliasi institusi. Selain rasionalitas harga dan kualitas, penelitian ini juga menemukan bahwa implementasi E-CRM (X3) memiliki peran yang signifikan dalam mendorong volume penjualan. Dalam konteks penjualan cat skala maklon (B2B), interaksi digital yang responsif, kemudahan pemesanan, serta tindak lanjut (follow-up) keluhan pelanggan melalui sistem elektronik terbukti efektif dalam menjaga loyalitas konsumen. Hal ini memperkuat gagasan bahwa kualitas produk yang baik harus ditunjang dengan sistem pelayanan dan retensi pelanggan yang terintegrasi secara digital.

Temuan kedua, sekaligus menjadi penyelesaian teknis dari penelitian ini, adalah pembuktian superioritas algoritma Random Forest atas model regresi klasik. Dalam banyak kasus regresi linier di dunia bisnis, keberadaan variabel Harga dan Kualitas seringkali berjalan beriringan (harga tinggi diasumsikan kualitas tinggi, dan sebaliknya). Korelasi ganda ini merusak asumsi klasik regresi linier (multikolinearitas) sehingga menaikkan angka kesalahan (RMSE 473 Kg). Sebaliknya, algoritma Random Forest, yang memecah dataset historis SMTI menjadi ratusan cabang pohon keputusan (decision trees) independen secara acak, berhasil mengabaikan kolinearitas tersebut dan mereduksi tingkat kesalahan prediksi secara dramatis menjadi 435,52 Kg. Angka deviasi ini sangat vital dalam skala industri, di mana penghematan kesalahan tebakan sebesar ±37 Kg per bulan dapat mencegah penumpukan aset mati di gudang yang membebani *cash flow* unit bisnis (Feizabadi, 2022).

Lebih lanjut, ketika model Random Forest ini digunakan sebagai simulator (mesin waktu) untuk memproyeksikan permintaan periode selanjutnya, hasil prediksi menunjukkan angka ekuilibrium permintaan berada di kisaran 3.025,59 Kg. Fakta analitik ini memberikan ketenangan manajerial, karena proyeksi tersebut masih berada jauh di bawah ambang batas kapasitas maksimum pabrik (5.000 Kg/bulan). Dengan demikian, kepala pabrik dapat menjadwalkan pembelanjaan bahan baku secara presisi tanpa perlu ketakutan akan terjadinya stockout. Inilah esensi utama dari transformasi menuju

data-driven decision making, di mana insting dikalibrasi oleh validitas kecerdasan buatan.

SIMPULAN

Penelitian ini berhasil membuktikan bahwa analitika prediktif menggunakan pendekatan Machine Learning jauh lebih superior dibandingkan penentuan kuota operasional berbasis insting di lingkungan produksi TEFA SMK SMTI. Kesimpulan utama dari riset ini menegaskan bahwa algoritma Random Forest Regression merupakan model peramalan terbaik (RMSE terendah sebesar 435,52 Kg) yang sangat tangguh dalam menangani data historis bervariasi dengan gejala multikolinearitas. Analisis inferensial juga membuktikan bahwa konsumen lebih mementingkan faktor rasionalitas harga, kualitas produk, serta pelayanan elektronik yang responsif (E-CRM), alih-alih citra sekolah (yang terbukti sebagai variabel tidak signifikan).

Saran yang dapat diberikan bagi manajemen TEFA SMTI Bandar Lampung adalah segera merevisi Standar Operasional Prosedur (SOP) perencanaan produksi dengan menjadikan model AI ini sebagai dashboard utama simulasi bulanan. Adapun untuk penelitian lanjutan yang dirasa perlu demi penyempurnaan hasil di masa depan, disarankan untuk mengintegrasikan variabel makro-ekonomi eksternal (seperti laju inflasi bahan baku kimia global atau tren daya beli masyarakat lokal) serta mencoba algoritma Deep Learning (Neural Network) dengan durasi data deret waktu yang lebih panjang (di atas 5 tahun) guna menangkap pola musiman (*seasonality*) yang lebih kompleks.

KONTRIBUSI

Penulis mengucapkan terima kasih dan penghormatan yang sebesar-besarnya kepada jajaran pimpinan serta staf ahli di unit bisnis Teaching Factory (TEFA) SMK SMTI Bandar Lampung atas keterbukaan, kolaborasi, dan izin pemanfaatan data operasional yang diberikan selama masa riset ini. Penghargaan juga ditujukan kepada program studi Magister Manajemen Teknologi, Institut Informatika dan Bisnis (IIB) Darmajaya atas bimbingan dan fasilitas akademik yang mendukung terselesaikannya penelitian ini.

REFERENSI

Aamer, A., Eka Yani, L. P., & Alan Priyatna, I. M. (2021). Data analytics in the supply chain management: Review of machine learning applications in demand forecasting. *Operations and Supply Chain Management: An International Journal*, *14*(1), 1–13. https://doi.org/10.31387/oscm0440281

Adnyana, I. G. K. S., Agustini, K., & Suartama, I. K. (2025). Analisis sistematis model dan implementasi teaching factory dalam pendidikan vokasi serta dampaknya terhadap kesiapan kerja siswa SMK. *Cetta: Jurnal Ilmu Pendidikan*, *8*(4), 326–332. https://doi.org/10.37329/cetta.v8i4.4714

Bougie, R., & Sekaran, U. (2019). *Research methods for business: A skill building approach* (8th ed.). Wiley.

Direktorat Sekolah Menengah Kejuruan. (2023). *Panduan teaching factory Sekolah Menengah Kejuruan*.

Feizabadi, J. (2022). Machine learning demand forecasting and supply chain performance. *International Journal of Logistics Research and Applications*, *25*(2), 119–142. https://doi.org/10.1080/13675567.2020.1803246

Gulo, B., & Machfud, S. (2026). Implementation of the Naive Bayes method for catering sales prediction at PT Negara Rasa Indonesia. *Multidisciplinary Indonesian Center Journal (MICJO)*, *3*(2), 2641–2648. https://doi.org/10.62567/micjo.v3i2.2374

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Prentice Hall.

Kadir, M. I., & Tricahyono, D. (2023). Analyzing Cyclops application acceptance in Telkomsel Pamasuka using IDT and TAM. *Jurnal Bisnis dan Pemasaran Digital*, *3*(1), 21–40. https://doi.org/10.35912/jbpd.v3i1.4498

Mai, H. G., Le, T. H., & Phung, T. D. N. (2026). Critical factors driving purchasing intention on TikTok live-stream platform: Insights from Vietnam. *International Journal of Procurement Management*, *25*(2), 194–220. https://doi.org/10.1504/IJPM.2026.151114

Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2024). *Introduction to time series analysis and forecasting* (3rd ed.). Wiley.

Nia, L., & Anggalia, W. (2021). Pengaruh kemudahan dan manfaat berbelanja online terhadap niat beli ulang dilihat dari sikap konsumen. *Jurnal Bina Bangsa Ekonomika*, *14*(1), 70–78. https://doi.org/10.46306/jbbe.v14i1.47

Sugiyono. (2022). *Metode penelitian: Kuantitatif, kualitatif, dan R&D* (Cet. ke-27). Alfabeta.

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2017). *Data mining: Practical machine learning tools and techniques* (4th ed.). Morgan Kaufmann.

Wulandari, W., Japarianto, E., & Tandijaya, T. N. B. (2022). Penerapan Technology Acceptance Model (TAM) terhadap perilaku konsumen mobile banking di Surabaya. *Jurnal Manajemen Pemasaran*, *16*(2), 126–132. https://doi.org/10.9744/pemasaran.16.2.126-132

Yamane, T. (1967). *Statistics: An introductory analysis*. Harper and Row.

Yufi, S. S., Fadli, U. M. D., & Khalida, L. R. (2025). Pengaruh Technology Acceptance Model (TAM) terhadap preferensi konsumen marketplace Shopee di kalangan karyawan Peruri. *COSTING: Journal of Economic, Business and Accounting*, *8*(4). https://doi.org/10.31539/6689tp27