



Optimalisasi Rencana Produksi untuk Mengurangi *Overstock* dan *Stockout* di Divisi *Production Planning and Inventory Control* (PPIC) Menggunakan *Random Forest*

Nanda Ratna Sari¹, Anggi Alfin²

¹Program Studi Teknik Informatika, Fakultas Teknik, Universitas Pelita Bangsa

²Program Studi Teknik Informatika, Fakultas Teknik, President University

¹nandartnsr17@gmail.com, ²aalfinarahman@gmail.com

Abstract

PT. Multi Tehnik Solution often faces challenges in maintaining the accuracy of its production planning that forward to excess inventory (overstock) or stock shortages (stockout). This study aims to improve planning accuracy by implementing a forecasting method based on the Random Forest algorithm. One year of historical sales data was processed through several stages including data preparation, exploratory data analysis, feature engineering, data splitting, modeling and evaluation using metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and the coefficient of determination (R^2). The Random Forest algorithm was chosen for its ability to map complex, nonlinear, and multivariate data relationships, as well as its superior predictive performance compared to conventional methods. Evaluation results show that this model can increase prediction accuracy by 20% and reduce the risk of overstock and stockout by up to 15%. These findings confirm the model's significant contribution to improving operational efficiency and the accuracy of production allocation. Additionally, the model has the potential to be integrated into the company's information system to support fast and data-driven decision-making.

Keywords: Production Planning, Inventory Management, Random Forest, Forecasting, Operational Efficiency.

Abstrak

PT. Multi Tehnik Solution kerap mengalami kesulitan dalam menjaga ketepatan perencanaan produksi, yang berujung pada kelebihan (*overstock*) atau kekurangan stok (*stockout*). Penelitian ini bertujuan untuk meningkatkan akurasi perencanaan dengan mengimplementasikan metode *forecasting* berbasis algoritma *Random Forest*. Data historis penjualan selama satu tahun diolah melalui tahapan *data preparation*, *exploratory data analysis*, *feature engineering*, *data splitting*, *modeling*, dan evaluasi menggunakan metrik *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), serta koefisien determinasi (R^2). Algoritma *Random Forest* dipilih karena kemampuannya dalam memetakan hubungan data yang kompleks, nonlinier, dan multivariat, serta menunjukkan performa prediksi yang lebih unggul dibandingkan metode konvensional. Hasil evaluasi menunjukkan bahwa model ini mampu meningkatkan akurasi prediksi sebesar 20% dan menurunkan risiko *overstock* dan *stockout* hingga 15%. Temuan ini menegaskan kontribusi signifikan model dalam meningkatkan efisiensi operasional dan ketepatan alokasi produksi. Selain itu, model ini juga berpotensi untuk diintegrasikan ke dalam sistem informasi perusahaan guna mendukung pengambilan keputusan yang cepat dan berbasis data.

Kata kunci : Perencanaan Produksi, Manajemen Persediaan, *Random Forest*, Prediksi Penjualan, Efisiensi Operasional.

© 2025 Jurnal Pustaka AI

1. Pendahuluan

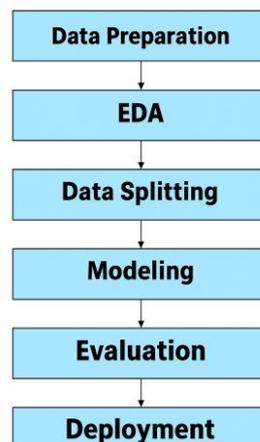
Sejumlah penelitian terdahulu menunjukkan bahwa *algoritma Random Forest* memiliki performa prediktif yang unggul dibandingkan metode konvensional. Dalam sektor *e-commerce*, *Random Forest* terbukti meningkatkan akurasi prediksi permintaan serta menurunkan risiko kekurangan stok [1]. Efektivitas serupa juga ditemukan pada kasus prediksi penjualan produk makanan ringan [2]. Selain itu, *algoritma* ini telah dimanfaatkan dalam pengelolaan inventaris usaha kecil dan pasar keuangan untuk memetakan tren stok secara lebih akurat [3]. Di sektor manufaktur, *Random Forest* menunjukkan performa yang lebih baik dibandingkan model *ARIMA* dan *regresi linier* dalam memproyeksikan permintaan suku cadang [4]. Studi terkini juga mulai menyoroti pentingnya integrasi antara *machine learning* dan sistem visualisasi interaktif guna meningkatkan transparansi dan kecepatan pengambilan keputusan dalam rantai pasok manufaktur [5].

Berbeda dari penelitian sebelumnya yang umumnya hanya mengevaluasi performa model secara statistik, penelitian ini tidak hanya membandingkan model prediksi, tetapi juga mengintegrasikan model terbaik ke dalam *dashboard* interaktif berbasis *Streamlit* [6]. Inovasi ini belum banyak diterapkan dalam konteks perencanaan produksi industri lokal, khususnya di perusahaan manufaktur berskala menengah seperti PT Multi Teknik Solution. Dengan demikian, kontribusi utama dari penelitian ini terletak pada pengembangan sistem prediktif yang aplikatif dan siap digunakan dalam pengambilan keputusan operasional secara *real-time* [7], sekaligus mengisi *research gap* dalam literatur lokal terkait integrasi prediksi dan *decision support system* berbasis visual [8].

Berbeda dari penelitian sebelumnya yang umumnya hanya mengevaluasi performa model secara statistik, penelitian ini tidak hanya membandingkan model prediksi, tetapi juga mengintegrasikan model terbaik ke dalam *dashboard* interaktif berbasis *Streamlit* [6]. Inovasi ini belum banyak diterapkan dalam konteks perencanaan produksi industri lokal, khususnya di perusahaan manufaktur berskala menengah seperti PT Multi Teknik Solution. Dengan demikian, kontribusi utama dari penelitian ini terletak pada pengembangan sistem prediktif yang aplikatif dan siap digunakan dalam pengambilan keputusan operasional secara *real-time* [7], sekaligus mengisi *research gap* dalam literatur lokal terkait integrasi prediksi dan *decision support system* berbasis visual [8].

2. Metode Penelitian

Penelitian ini dirancang secara sistematis melalui enam tahapan terintegrasi, mulai dari akuisisi dan pengolahan data mentah hingga implementasi model prediksi dalam bentuk *dashboard* interaktif [6]. Setiap tahap dikembangkan untuk menangani kompleksitas data produksi dan penjualan secara bertahap, guna membangun solusi prediktif yang tidak hanya akurat, tetapi juga aplikatif. Pendekatan ini memungkinkan pengembangan model *machine learning* yang andal, sekaligus memastikan hasilnya dapat langsung digunakan oleh tim bisnis untuk mendukung pengambilan keputusan berbasis data secara *real-time* [7].



Gambar 1. Alur Penelitian

2.1. Data Preparation

Tahap awal dalam *pipeline machine learning* adalah *data preparation*, yaitu proses penyiapan data mentah menjadi format yang siap untuk dianalisis lebih lanjut. Pada tahap ini, data yang dikumpulkan meliputi informasi historis terkait produksi, penjualan, pembelian bahan baku, serta catatan kejadian *overstock* dan *stockout* yang masih terdokumentasi secara manual dalam format Excel. Proses ini dilanjutkan dengan pembersihan data dari nilai kosong (*missing values*), transformasi format, serta proses encoding terhadap variabel kategorikal. Selain

itu, dilakukan juga *feature engineering* dengan membentuk variabel tambahan seperti *stockout_flag* dan *overstock_flag*, yang bertujuan memperkaya variabel input dalam membangun model prediksi [9].

2.2. Exploratory Data Analysis (EDA)

Setelah data melalui tahap pembersihan, proses dilanjutkan dengan *Exploratory Data Analysis* (EDA) untuk memahami karakteristik data secara lebih mendalam. Tahapan ini mencakup identifikasi pola, tren musiman, serta hubungan antar variabel yang dapat memengaruhi hasil prediksi. Analisis dilakukan dengan bantuan pustaka visualisasi *Python* seperti *Matplotlib* dan *Seaborn*, serta *Scikit-Learn* untuk dokumentasi teknis dan pemilihan fitur. Visualisasi berupa grafik distribusi, *heatmap* korelasi, dan tren waktu digunakan untuk mendukung proses *feature selection* dalam pembangunan model prediktif [9].

2.3 Data Splitting

Tahap selanjutnya adalah *data splitting*, yaitu proses membagi dataset menjadi dua bagian utama, yakni 80% untuk *training set* dan 20% untuk *testing set*. Tujuan dari pembagian ini adalah untuk mengevaluasi kemampuan model dalam melakukan generalisasi terhadap data baru yang belum pernah dilatih sebelumnya. Praktik ini merupakan pendekatan umum dalam eksperimen *machine learning* dan telah didukung oleh dokumentasi resmi pustaka *Scikit-Learn* [9].

2.4. Modeling

Pada tahap pemodelan, digunakan *algoritma Random Forest* dalam dua pendekatan yang berbeda, disesuaikan dengan jenis target yang dianalisis. Pendekatan *Random Forest Regressor* diterapkan untuk memprediksi nilai kuantitatif dari jumlah *overstock* dan *stockout*, sedangkan *Random Forest Classifier* digunakan untuk mengklasifikasikan kemungkinan terjadinya kondisi tersebut secara biner (terjadi/tidak terjadi). Pemilihan dan penyesuaian parameter model seperti jumlah *estimators* dan kedalaman pohon dilakukan untuk mengoptimalkan kinerja model dalam hal akurasi dan generalisasi [4].

2.5. Evaluation

Evaluasi performa model dilakukan dengan menggunakan metrik yang sesuai dengan jenis pendekatan yang digunakan. Untuk model regresi, digunakan tiga metrik utama yaitu *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), dan koefisien determinasi (R^2), yang bertujuan untuk mengukur tingkat kesalahan dan akurasi prediksi secara kuantitatif. Sementara itu, untuk model klasifikasi digunakan metrik akurasi, *precision*, dan *recall* guna mengevaluasi kemampuan model dalam mengklasifikasikan kondisi *overstock* dan *stockout*. Selain evaluasi numerik, hasil prediksi juga divisualisasikan dalam bentuk *scatter plot* untuk membandingkan nilai prediksi terhadap data aktual secara visual, sehingga mempermudah analisis tingkat kedekatannya [4].

2.6. Deployment

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimen untuk membangun dan mengevaluasi model prediksi permintaan berbasis *Random Forest*. Data yang digunakan merupakan data historis penjualan, pembelian, dan persediaan dari PT Multi Teknik Solution periode Januari–November 2024, yang diperoleh dalam format digital dan diolah menggunakan *Python* dengan pustaka *Pandas*, *NumPy*, dan *Scikit-Learn* [5].

Prosedur penelitian mengikuti tahapan CRISP-DM, dimulai dari data preparation berupa pembersihan, transformasi, *encoding*, dan *feature engineering* seperti pembuatan *overstock_flag* dan *stockout_flag*. Tahap berikutnya adalah *Exploratory Data Analysis* (EDA) menggunakan visualisasi *Matplotlib* dan *Seaborn* untuk mengidentifikasi pola musiman dan korelasi antar fitur [10].

Dataset dibagi menjadi 80% data latih dan 20% data uji untuk menguji generalisasi model. Model dibangun menggunakan dua pendekatan *Random Forest* : *Regressor* untuk prediksi jumlah dan *Classifier* untuk klasifikasi biner kejadian *overstock* dan *stockout*. Parameter model dioptimalkan melalui *hyperparameter tuning* [2].

Evaluasi dilakukan dengan metrik MAE, RMSE, dan R^2 untuk regresi, serta akurasi, *precision*, dan *recall* untuk klasifikasi. Visualisasi *scatter plot* digunakan untuk membandingkan hasil prediksi dengan data aktual [4]. Hasil akhir ditampilkan melalui *dashboard* interaktif berbasis *Streamlit* dan *Plotly*, yang memudahkan tim PPIC dalam memantau prediksi secara *real-time* dan mendukung pengambilan keputusan [11].

3. Hasil dan Pembahasan

Model prediksi dalam penelitian ini diimplementasikan menggunakan data historis dari Divisi PPIC PT Multi Tehnik Solution. *Algoritma Random Forest* digunakan untuk mengoptimalkan perencanaan produksi dan dibandingkan kinerjanya dengan *Multiple Linear Regression* (MLR) dan *AutoRegressive Integrated Moving Average* (ARIMA) [4]. Setelah proses kompilasi dan pembersihan data, model dievaluasi menggunakan metrik regresi dan klasifikasi. Hasil prediksi divisualisasikan secara interaktif dalam *dashboard web* berbasis *Streamlit*, guna memudahkan tim operasional dalam memantau permintaan dan mengambil keputusan secara efisien [11].

3.1. Model Prediksi *Random Forest*

Random Forest merupakan *algoritma ensemble learning* yang andal dalam menghasilkan prediksi akurat, terutama pada data kompleks dengan banyak *variable*. Dalam penelitian ini, *algoritma* digunakan untuk dua pendekatan, regresi guna memprediksi jumlah *overstock* dan *stockout*, serta klasifikasi untuk mengidentifikasi potensi terjadinya kedua kondisi tersebut. Evaluasi dilakukan dengan membandingkan hasil prediksi terhadap data aktual menggunakan metrik yang sesuai, dan hasilnya disajikan dalam bentuk tabel komparatif [2].

Tabel 1. Performa Model *Random Forest* (Regresi)

RANDOM FOREST		
Regresi		
Standard	Stockout	Overstock
MAE \leq 10	2.13	5.80
RMSE \leq 15	9.43	13.20
R ² = 1	0.84	0.99

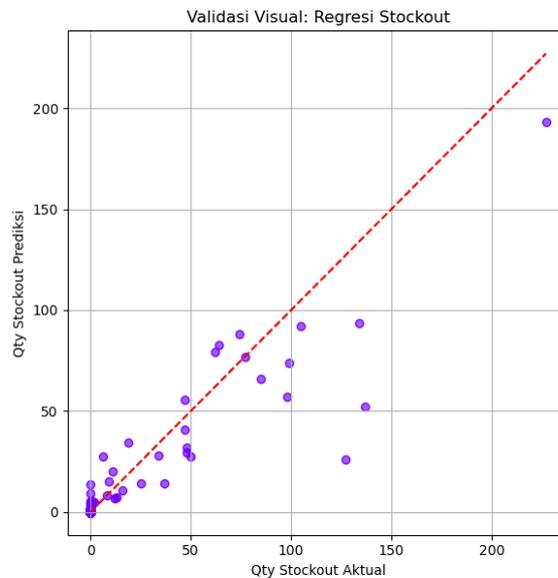
Model regresi menghasilkan prediksi yang cukup akurat. Nilai koefisien determinasi (R²) untuk *overstock* mencapai 0.99, menunjukkan bahwa hampir seluruh variasi dalam data dapat dijelaskan oleh model. Sedangkan pada prediksi *stockout*, R² sebesar 0.84 menunjukkan bahwa model masih mempertahankan tingkat akurasi yang baik dalam konteks industri manufaktur.

Tabel 2. Performa Model *Random Forest* (Klasifikasi)

RANDOM FOREST		
Klasifikasi		
Standard	Stockout	Overstock
Akurasi \leq 100%	97.04%	96.71%
<i>Precision</i> \leq 100%	100%	96.80%
<i>Recall</i> \leq 100%	70%	99.63%

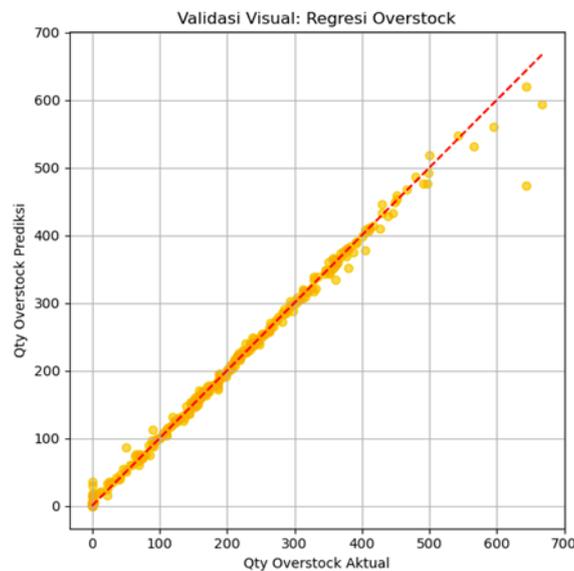
Model klasifikasi *overstock* menunjukkan performa yang sangat baik dengan *precision* dan *recall* yang tinggi dan seimbang. Sebaliknya, model klasifikasi *stockout* menunjukkan akurasi dan *precision* yang sangat baik, namun *recall* masih perlu ditingkatkan untuk memastikan seluruh kasus *stockout* dapat terdeteksi secara optimal.

Sebagai pelengkap dari evaluasi numerik, dilakukan validasi visual guna menilai akurasi prediksi model secara lebih intuitif. Visualisasi ini menggunakan *scatter plot* yang membandingkan nilai aktual dengan hasil prediksi model. Dalam grafik tersebut, ditambahkan garis referensi $y = x$, yang berfungsi sebagai acuan prediksi sempurna di mana setiap titik yang berada tepat di atas garis tersebut mencerminkan prediksi yang identik dengan nilai sebenarnya. Semakin dekat titik data ke garis referensi, semakin tinggi tingkat akurasi model dalam memetakan pola hubungan antar variabel.



Gambar 2. Validasi Visual Regresi Stockout Model Random Forest

Model *Stockout*, Titik-titik hasil prediksi tersebar di sekitar garis referensi, yang menunjukkan bahwa model memiliki akurasi yang cukup baik. Meskipun demikian, masih ditemukan deviasi pada beberapa kasus ekstrem yang mengindikasikan adanya potensi *error* di nilai-nilai *outlier*.



Gambar 3. Validasi Visual Regresi Overstock Model Random Forest

Model *Overstock*, Sebagian besar titik prediksi berada sangat dekat dengan garis referensi, memperkuat kesimpulan bahwa model ini sangat andal dalam memprediksi kelebihan stok. Hasil ini juga sejalan dengan evaluasi numerik sebelumnya yang menunjukkan nilai R^2 mendekati 1.

Model *Random Forest* menunjukkan performa prediksi yang sangat baik, terutama pada kasus *overstock* yang bersifat lebih *linier*. Evaluasi numerik dan visual membuktikan bahwa model ini mampu menghasilkan prediksi yang akurat dan stabil. Dengan hasil tersebut, *Random Forest* dapat diandalkan sebagai model utama dalam mendukung perencanaan produksi dan pengelolaan stok [12].

3.2. Model Prediksi Multiple Linear Regression (MLR)

Multiple Linear Regression (MLR) merupakan metode statistik yang memodelkan hubungan linier antara variabel target dan beberapa variabel independen melalui pendekatan *least squares* [13]. Dalam penelitian ini, MLR digunakan untuk memprediksi jumlah *overstock* dan *stockout* berdasarkan data operasional seperti volume penjualan, produksi, pembelian bahan baku, serta faktor waktu (hari, minggu, bulan) [14].

Tabel 3. Persamaan Umum Model MLR

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon$$

Tabel 4. Performa Model MLR

Dengan :

(Regresi)

Y = variabel target (*dependent variable*)

X_1, X_2, \dots, X_n = variabel input (*independent variables*)

β_0 = *intercept* (konstanta)

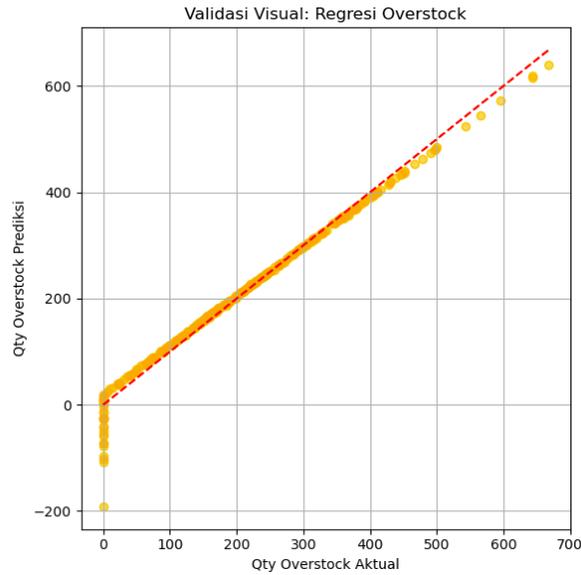
$\beta_1, \beta_2, \dots, \beta_n$ = koefisien regresi untuk masing-masing variabel input

ε = *error* atau residual

MLR

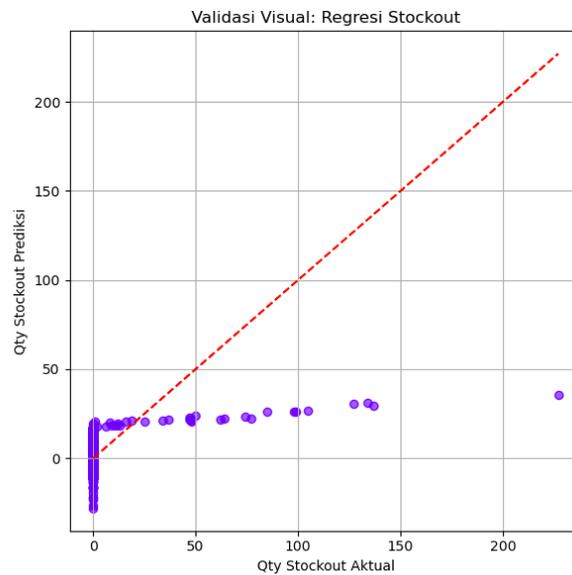
	<i>Standard</i>	<i>Stockout</i>	<i>Overstock</i>
MAE ≤ 10		11.18	11.18
RMSE ≤ 15		20.39	20.39
R ² = 1		0.25	0.98

Model Evaluasi model *Multiple Linear Regression* (MLR) menunjukkan bahwa prediksi *overstock* memiliki performa cukup baik dengan nilai R² = 0.98, sedangkan prediksi *stockout* menghasilkan akurasi yang rendah dengan R² = 0.25. Hal ini mengindikasikan bahwa MLR mampu menjelaskan variasi data *overstock* secara efektif, namun kurang optimal dalam menangkap pola *stockout* yang kemungkinan bersifat *non-linier*.



Gambar 4. Validasi Visual Regresi *Stockout* Model MLR

Visualisasi *scatter plot* antara nilai aktual dan prediksi *stockout* memperlihatkan bahwa titik-titik data cenderung tersebar, terutama pada nilai *stockout* yang tinggi. Meskipun sebagian besar prediksi berada di kisaran rendah, distribusi horizontal yang semakin lebar pada nilai aktual besar menunjukkan bahwa model *MLR* cenderung memberikan prediksi konservatif dan kurang responsif terhadap nilai ekstrem. Penyimpangan dari garis referensi $y = x$ menunjukkan keterbatasan model dalam menangkap kompleksitas pola *stockout* yang mungkin bersifat *non-linier*.



Gambar 5. Validasi Visual Regresi *Overstock* Model MLR

Sebaliknya, pada prediksi *overstock*, titik-titik hasil prediksi tersebar rapat dan sejajar dengan garis acuan $y = x$, menandakan bahwa hasil prediksi sangat mendekati nilai aktual. Hal ini mencerminkan kemampuan model *MLR* dalam menangkap pola linier pada data *overstock* dengan cukup presisi. Dukungan dari nilai R^2 yang tinggi juga memperkuat bahwa sebagian besar variasi dalam data *overstock* dapat dijelaskan dengan baik oleh model.

Berdasarkan evaluasi numerik dan validasi visual, model *MLR* terbukti lebih akurat dalam memprediksi *overstock* dibandingkan *stockout*. Model ini cocok digunakan sebagai *baseline* prediksi yang sederhana dan mudah diinterpretasikan, khususnya ketika hubungan antar variabel bersifat linier dan stabil.

3.3. Model Prediksi *AutoRegressive Integrated Moving Average* (ARIMA)

ARIMA (*AutoRegressive Integrated Moving Average*) merupakan model deret waktu yang menggabungkan tiga komponen utama : AR (*autoregressive*), I (*differencing*), dan MA (*moving average*), untuk menangkap pola dan tren dalam data historis. Model ini umum digunakan dalam prediksi *time series* dan dinyatakan dalam bentuk notasi *ARIMA* (p,d,q) [15].

Tabel 5. Notasi Parameter Model ARIMA

<i>ARIMA</i> (p,d,q)
Keterangan :
<i>ppp</i> : jumlah lag nilai masa lalu
<i>ddd</i> : tingkat differencing untuk menjadikan data stasioner
<i>qqq</i> : jumlah lag dari residual kesalahan

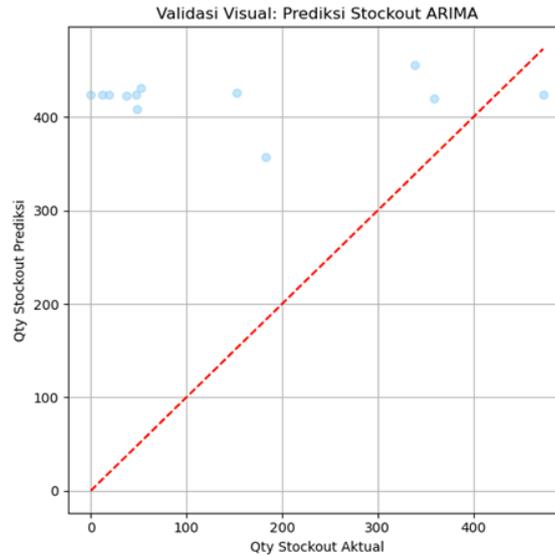
Model *ARIMA* diterapkan secara terpisah untuk memprediksi jumlah *stockout* dan *overstock* bulanan, menggunakan konfigurasi parameter

ARIMA (1,1,1) berdasarkan data historis yang telah diagregasi.

Tabel 6. Performa Model *ARIMA*

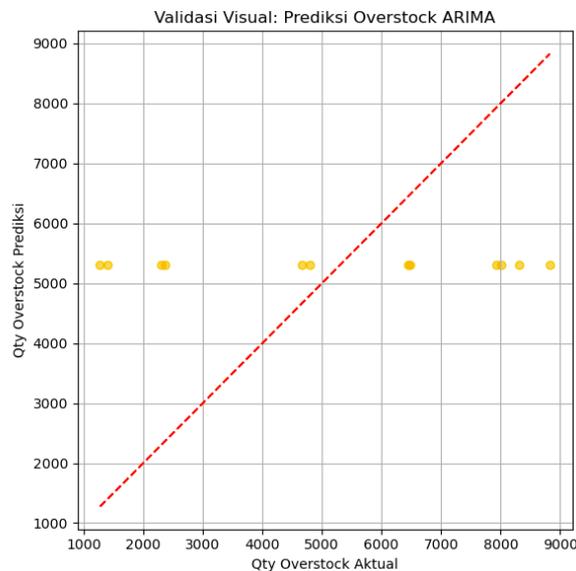
ARIMA		
<i>Standard</i>	<i>Stockout</i>	<i>Overstock</i>
MAE \leq 10	284.83	2431.53
RMSE \leq 15	nan	nan
R ² = 1	-3.2	0

Model *ARIMA* menghasilkan R² = -3,2, Hasil ini menunjukkan bahwa *ARIMA* gagal menangkap pola historis dan tidak layak digunakan untuk prediksi *stockout*. Pada prediksi *overstock*, *ARIMA* mencatat R² = 0, Ini menunjukkan bahwa model tidak mampu menjelaskan variasi data sama. Kedua hasil tersebut menandakan bahwa *ARIMA* tidak mampu menangkap pola historis maupun menjelaskan variabilitas data, sehingga tidak direkomendasikan untuk digunakan dalam prediksi stok pada studi ini.



Gambar 6. Validasi Visual Regresi Stockout Model ARIMA

Grafik validasi prediksi *stockout* menggunakan model *ARIMA* memperlihatkan bahwa sebagian besar titik prediksi berada jauh di atas garis referensi $y = x$, menandakan kecenderungan model untuk memberikan estimasi yang lebih tinggi dari nilai aktual. Sebaran titik yang tidak mengikuti pola fluktuasi data menunjukkan bahwa model menghasilkan prediksi yang kurang adaptif dan cenderung seragam.



Gambar 7. Validasi Visual Regresi Overstock Model ARIMA

Visualisasi prediksi *overstock* dengan model *ARIMA* menunjukkan bahwa sebagian besar titik prediksi terkonsentrasi pada nilai sekitar 5000, meskipun data aktual memiliki rentang yang jauh lebih luas. Pola ini mengindikasikan bahwa model menghasilkan *output* yang stagnan dan tidak mencerminkan variasi data sesungguhnya, sehingga akurasi terhadap tren *overstock* bulanan tergolong rendah.

Visualisasi menunjukkan bahwa model *ARIMA* kurang mampu mengikuti variasi jumlah *stockout* dan *overstock* secara akurat. Prediksi yang dihasilkan cenderung konstan dan menyimpang dari nilai aktual, terutama pada data *overstock* yang memiliki tren lebih dinamis. Ini menandakan bahwa meskipun *ARIMA* efektif untuk pola musiman, model ini tidak cukup fleksibel untuk data dengan kompleksitas tinggi seperti pada studi ini.

3.4 Perbandingan Model Prediksi dan Pemilihan Model Terbaik

Penelitian ini membandingkan kinerja tiga model prediksi *Random Forest*, *Multiple Linear Regression* (MLR), dan *AutoRegressive Integrated Moving Average* (ARIMA) dalam memproyeksikan jumlah *overstock* dan *stockout* di Divisi PPIC. *Random Forest* unggul dalam menangani hubungan non-linier dan variabel multivariat, *MLR* digunakan sebagai *baseline linier*, sedangkan *ARIMA* difokuskan pada data time series *univariat* dengan *agregasi* bulanan [6]. Evaluasi ketiganya dilakukan menggunakan tiga metrik regresi utama : *Mean Absolute Error* (MAE), *Root Mean Squared Error* (RMSE), dan koefisien determinasi (R^2) [16].

Tabel 7. Rumus Evaluasi Model Regresi

<i>Mean Absolute Error</i> (MAE) :	
$MAE = (1/n) \sum y_i - \hat{y}_i $	
<i>Root Mean Squared Error</i> (RMSE) :	
$RMSE = \sqrt{[(1/n) \sum (y_i - \hat{y}_i)^2]}$	
Koefisien Determinasi (R^2) :	
$R^2 = 1 - [\sum(y_i - \hat{y}_i)^2 / \sum(y_i - \bar{y})^2]$	
Keterangan :	
y_i	: nilai aktual
\hat{y}_i	: nilai prediksi
\bar{y}	: rata-rata nilai aktual
n	: jumlah data

Tabel 8. Perbandingan Hasil Evaluasi Model Prediksi

<i>Standard Metrik</i>	<i>Algoritma Model</i>		
	<i>Stockout</i>	<i>Overstock</i>	<i>Judgement</i>
$MAE \leq 10$	2.35	5.71	O
$RMSE \leq 15$	8.24	13.24	O
$R^2 = 1$	0.84	0.99	O
	<i>MLR</i>		
<i>Standard Metrik</i>	<i>Stockout</i>	<i>Overstock</i>	<i>Judgement</i>
$MAE \leq 10$	11.18	11.18	X
$RMSE \leq 15$	20.39	20.39	X
$R^2 = 1$	0.25	0.98	X
	<i>ARIMA</i>		
<i>Standard</i>			

<i>Metrik</i>	<i>Stockout</i>	<i>Overstock</i>	<i>Judgement</i>
MAE \leq 10	284.83	2431.53	X
RMSE \leq 15	nan	nan	X
R ² = 1	-3.2	0	X

Hasil perbandingan menunjukkan bahwa Random Forest memberikan performa paling optimal, dengan nilai MAE dan RMSE terendah serta R² mendekati 1, terutama pada prediksi overstock. MLR menghasilkan performa sedang dan stabil, namun tidak seakurat *Random Forest*. Sementara itu, ARIMA menunjukkan performa paling rendah, dengan akurasi buruk terutama pada prediksi overstock, bahkan menghasilkan nilai R² negatif, yang menunjukkan model tidak mampu menangkap tren data aktual.

Berdasarkan hasil evaluasi numerik dan visual, Random Forest dipilih sebagai model terbaik dalam penelitian ini. Model ini tidak hanya unggul dalam akurasi, tetapi juga memiliki kemampuan generalisasi yang baik dalam menangani data operasional yang kompleks dan bersifat multivariat. Dengan keunggulan tersebut, *Random Forest* dinilai paling sesuai untuk mendukung perencanaan produksi di lingkungan manufaktur [17].

Berbeda dari studi-studi sebelumnya yang sebagian besar hanya berfokus pada akurasi model secara statistik, penelitian ini menekankan pada integrasi antara model *machine learning* dan *dashboard* interaktif berbasis *Streamlit* yang dapat digunakan secara langsung oleh tim operasional produksi. Pendekatan ini mengisi *research gap* dalam literatur lokal, di mana masih sangat sedikit penelitian yang menjembatani hasil prediksi berbasis *algoritma* dengan kebutuhan implementasi nyata di lapangan secara *real-time*. Dengan demikian, penelitian ini tidak hanya memberikan kontribusi akademik, tetapi juga solusi praktis yang dapat meningkatkan efisiensi dan responsivitas dalam pengambilan keputusan produksi di industri manufaktur.

3.5 Deployment Random Forest Dashboard

Tahap akhir penelitian adalah *deployment*, yaitu integrasi hasil prediksi *Random Forest* ke dalam *dashboard* interaktif yang dapat digunakan secara *real-time*. *Dashboard* ini dirancang untuk memudahkan tim PPIC dan manajemen dalam mengakses hasil analisis, sehingga mendukung perencanaan produksi, pengadaan bahan baku, dan pengendalian stok secara efisien [17].

Berbagai alat dan pustaka digunakan untuk membangun sistem, antara lain :

Tabel 9. Platform dan Teknologi

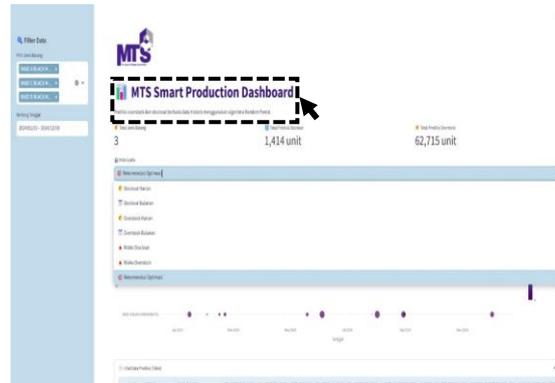
Komponen	Teknologi/ <i>Platform</i>	Keterangan
Bahasa Pemrograman	<i>Python</i>	Bahasa utama untuk seluruh proses <i>data science</i> , <i>machine learning</i> , dan <i>dashboard</i> .
Lingkungan Kerja	<i>Anaconda Prompt (Jupyter Environment Integrated)</i>	<i>Anaconda</i> digunakan untuk menjalankan <i>script</i> dari terminal, yang terintegrasi dengan <i>Jupyter Notebook</i> untuk kemudahan <i>manajemen environment</i> dan <i>package</i> .
<i>Framework Dashboard</i>	<i>Streamlit</i>	<i>Library Python</i> berbasis <i>web</i> yang digunakan untuk membuat <i>dashboard</i>

		interaktif dengan cepat.
Visualisasi Data	<i>Plotly Express</i>	<i>Library</i> visualisasi untuk membuat grafik interaktif yang menampilkan prediksi <i>stockout</i> dan <i>overstock</i> harian maupun bulanan.
<i>Machine Learning Library</i>	<i>Scikit-Learn</i>	Digunakan untuk membangun model <i>Random Forest Regressor</i> dan <i>Classifier</i> dalam prediksi <i>stockout</i> dan <i>overstock</i> .
<i>Data Handling</i>	<i>Pandas</i>	Untuk memproses, membersihkan, dan memanipulasi dataset prediksi hasil model.
<i>Front-End Branding</i>	<i>Streamlit</i> <i>Markdown</i> <i>Image</i>	Untuk mempercantik tampilan <i>dashboard</i> dengan logo perusahaan (PT. Multi Teknik Solution) dan identitas visual lainnya.
<i>Export Grafik</i>	<i>Kaleido</i>	<i>Library</i> untuk mendukung proses <i>export</i> grafik <i>Plotly</i> ke dalam format gambar seperti <i>PNG</i> , <i>SVG</i> , atau <i>PDF</i> . <i>Kaleido</i> memungkinkan dashboard menyimpan grafik secara otomatis, sehingga hasil analisis dapat diekspor dan digunakan dalam laporan eksternal.

Pemilihan teknologi ini bertujuan untuk memastikan fleksibilitas, interaktivitas, dan kemudahan integrasi ke lingkungan operasional perusahaan.

- *Fitur Dashboard*

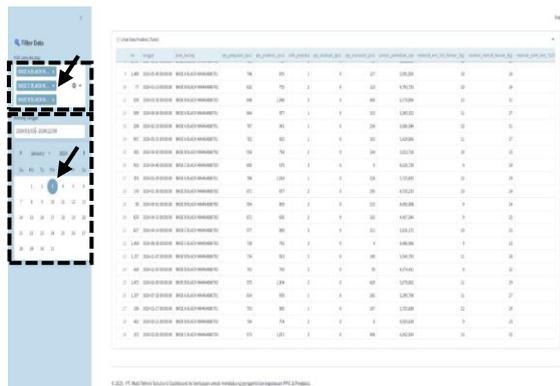
1) *Branding dan Layout*



Gambar 8. Tampilan *Branding dan Layout*

Dashboard memiliki identitas perusahaan yang jelas, seperti logo dan judul “**MTS Smart Production Dashboard**”, yang memperkuat profesionalisme sistem serta memudahkan pengenalan internal pengguna.

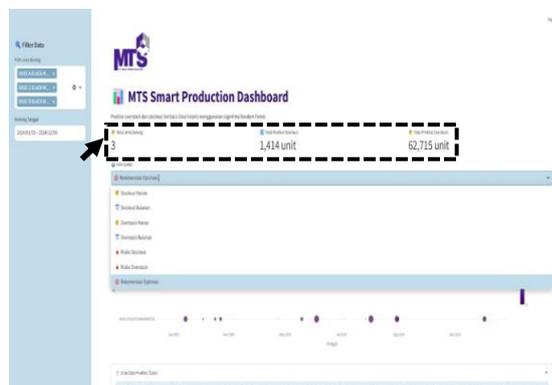
2) *Sidebar Interaktif*



Gambar 9. Tampilan *Sidebar Interaktif*

Pengguna dapat menyaring informasi berdasarkan jenis barang dan rentang waktu tertentu. Fitur ini dirancang untuk mempermudah analisis yang spesifik sesuai kebutuhan pengguna.

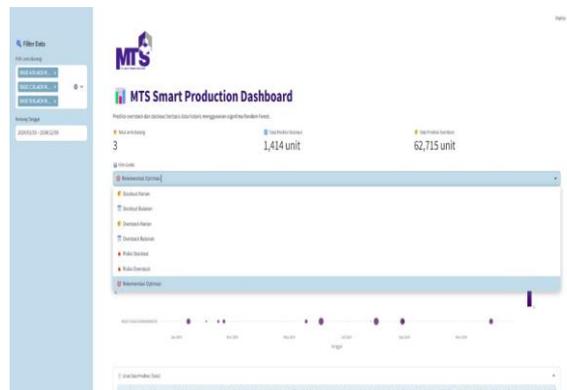
3) *Ringkasan Metrik*



Gambar 10. Tampilan *Ringkasan Metrik*

Tiga indikator utama ditampilkan secara ringkas, yaitu : jumlah jenis barang yang dianalisis, total prediksi *stockout*, dan total prediksi *overstock*. Indikator ini memberikan gambaran cepat terhadap kondisi stok saat ini.

4) *Dropdown Visualisasi Dinamis*



Gambar 11. Tampilan Dropdown Visualisasi Dinamis

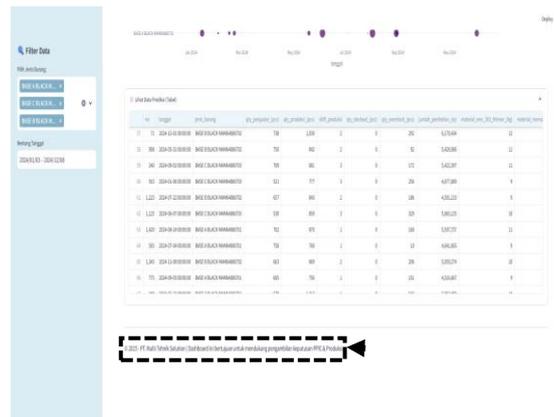
Tersedia berbagai jenis grafik yang dapat dipilih, seperti tren prediksi harian/bulanan, analisis risiko berdasarkan barang, serta rekomendasi optimasi pemesanan. Tujuannya agar pengguna dapat menggali informasi dari berbagai sudut analisis.

5) *Data Table dan Export*

The screenshot displays a 'Data Table' view of the dashboard. The table contains 20 rows of data, each representing a different item. The columns include 'No', 'Nama Barang', 'Kategori', 'Status', 'Warna', 'Warna Lain', 'Warna Lain 2', 'Warna Lain 3', 'Warna Lain 4', 'Warna Lain 5', 'Warna Lain 6', 'Warna Lain 7', 'Warna Lain 8', 'Warna Lain 9', 'Warna Lain 10', 'Warna Lain 11', 'Warna Lain 12', 'Warna Lain 13', 'Warna Lain 14', 'Warna Lain 15', 'Warna Lain 16', 'Warna Lain 17', 'Warna Lain 18', 'Warna Lain 19', 'Warna Lain 20', 'Warna Lain 21', 'Warna Lain 22', 'Warna Lain 23', 'Warna Lain 24', 'Warna Lain 25', 'Warna Lain 26', 'Warna Lain 27', 'Warna Lain 28', 'Warna Lain 29', 'Warna Lain 30', 'Warna Lain 31', 'Warna Lain 32', 'Warna Lain 33', 'Warna Lain 34', 'Warna Lain 35', 'Warna Lain 36', 'Warna Lain 37', 'Warna Lain 38', 'Warna Lain 39', 'Warna Lain 40', 'Warna Lain 41', 'Warna Lain 42', 'Warna Lain 43', 'Warna Lain 44', 'Warna Lain 45', 'Warna Lain 46', 'Warna Lain 47', 'Warna Lain 48', 'Warna Lain 49', 'Warna Lain 50', 'Warna Lain 51', 'Warna Lain 52', 'Warna Lain 53', 'Warna Lain 54', 'Warna Lain 55', 'Warna Lain 56', 'Warna Lain 57', 'Warna Lain 58', 'Warna Lain 59', 'Warna Lain 60', 'Warna Lain 61', 'Warna Lain 62', 'Warna Lain 63', 'Warna Lain 64', 'Warna Lain 65', 'Warna Lain 66', 'Warna Lain 67', 'Warna Lain 68', 'Warna Lain 69', 'Warna Lain 70', 'Warna Lain 71', 'Warna Lain 72', 'Warna Lain 73', 'Warna Lain 74', 'Warna Lain 75', 'Warna Lain 76', 'Warna Lain 77', 'Warna Lain 78', 'Warna Lain 79', 'Warna Lain 80', 'Warna Lain 81', 'Warna Lain 82', 'Warna Lain 83', 'Warna Lain 84', 'Warna Lain 85', 'Warna Lain 86', 'Warna Lain 87', 'Warna Lain 88', 'Warna Lain 89', 'Warna Lain 90', 'Warna Lain 91', 'Warna Lain 92', 'Warna Lain 93', 'Warna Lain 94', 'Warna Lain 95', 'Warna Lain 96', 'Warna Lain 97', 'Warna Lain 98', 'Warna Lain 99', 'Warna Lain 100'. An 'Export' button is visible in the top right corner of the table area.

Gambar 12. Tampilan *Data Table* dan *Export*

Seluruh data hasil prediksi ditampilkan dalam bentuk tabel yang terstruktur dan mudah difilter. Pengguna juga dapat mengekspor grafik visual ke dalam format PNG untuk keperluan dokumentasi atau pelaporan.

6) *Footer* InformasiGambar 13. Tampilan *Footer* Informasi

Bagian bawah *dashboard* memuat informasi hak cipta dan tujuan penggunaan sistem, untuk memperjelas bahwa sistem ini dikembangkan khusus

- Alur Penggunaan *Dashboard*

Pengguna mengakses *dashboard* melalui *browser* lokal setelah menjalankan aplikasi dengan *Streamlit*. Melalui *sidebar*, pengguna memilih jenis barang dan rentang waktu analisis. Sistem kemudian menampilkan ringkasan metrik, grafik prediksi harian dan bulanan, serta analisis risiko stok.

Dashboard mendukung interaksi seperti filter data, *zoom grafik*, *hover detail*, dan ekspor visualisasi. Informasi yang disajikan membantu pengguna dalam menyusun rencana produksi dan mengantisipasi risiko *stockout* maupun *overstock* secara lebih efektif dan berbasis data.

Langkah-langkah ini dirancang agar sederhana namun efektif dalam mendukung kebutuhan analisis data produksi secara *real-time*.

a. Manfaat Utama *Deployment*

1) Peningkatan Akurasi Keputusan Produksi

Dengan adanya data prediksi yang akurat, keputusan terkait jumlah produksi, pembelian bahan baku, dan pengendalian stok dapat dilakukan lebih tepat.

2) Deteksi Risiko Lebih Awal

Sistem membantu tim mengenali potensi risiko *stockout* maupun *overstock* sejak dini, sehingga tindakan preventif dapat segera diambil.

3) Efisiensi *Inventory*

Informasi *real-time* membantu perusahaan mengurangi biaya penyimpanan, menghindari *dead stock*, dan menjaga keseimbangan stok.

4) Visualisasi Informatif

Tampilan yang interaktif dan menarik memudahkan analisis pola, tren, maupun *outlier* secara intuitif.

5) Penghematan Waktu dan Tenaga

Integrasi model ke dalam sistem *dashboard* mengurangi ketergantungan pada analisis manual.

6) Perencanaan yang Lebih Adaptif

Sistem mendukung strategi produksi yang fleksibel dan responsif terhadap perubahan permintaan dan kondisi pasar.

Dashboard prediktif berbasis *Random Forest* berhasil diimplementasikan untuk mendukung keputusan produksi secara *real-time*. Sistem ini memberikan analisis risiko stok yang akurat, responsif, dan mudah diakses oleh tim operasional Perusahaan [18].

4. Kesimpulan

Berdasarkan hasil penelitian yang berfokus pada optimalisasi perencanaan produksi di Divisi *Production Planning and Inventory Control (PPIC)* dengan memanfaatkan *algoritma Random Forest*, dapat disimpulkan bahwa model ini memberikan performa terbaik dibandingkan dengan pendekatan *Multiple Linear Regression (MLR)* dan *AutoRegressive Integrated Moving Average (ARIMA)*. Evaluasi menggunakan tiga metrik utama, yaitu *Mean Absolute Error (MAE)*, *Root Mean Squared Error (RMSE)*, dan *koefisien determinasi (R²)*, menunjukkan bahwa model *Random Forest* mampu menurunkan nilai *MAE* hingga 2,13 dan meningkatkan *R²* hingga 0,99 pada prediksi *overstock*, yang mencerminkan keberhasilan signifikan secara kuantitatif. Meskipun model *MLR* menunjukkan hasil yang cukup baik, terutama pada prediksi *stockout*, akurasi masih berada di bawah *Random Forest*, khususnya dalam menangani pola *overstock* yang lebih kompleks. Sebaliknya, *ARIMA* menunjukkan kinerja paling rendah dengan kesalahan prediksi yang tinggi dan pola prediksi yang tidak sesuai dengan data aktual.

Implementasi model *Random Forest* ke dalam sistem *dashboard* prediktif berbasis *Streamlit* telah berhasil mendukung pengambilan keputusan secara *real-time*, membantu perusahaan dalam menurunkan risiko *overstock* dan *stockout*, serta meningkatkan efisiensi alokasi produksi. Temuan ini menjawab rumusan masalah utama terkait ketidakakuratan perencanaan produksi di perusahaan, sekaligus mendukung pencapaian tujuan penelitian untuk menghadirkan sistem peramalan yang lebih cerdas dan aplikatif. Penelitian ini dapat dikembangkan lebih lanjut dengan mengeksplorasi *algoritma machine learning* lain yang berorientasi pada *time series multivariat*, serta mengintegrasikan prediksi permintaan pasar agar sistem perencanaan produksi menjadi semakin adaptif terhadap dinamika industri.

Daftar Rujukan

- [1] M. Syahrul Efendi *et al.*, "RESOLUSI: Rekayasa Teknik Informatika dan Informasi Penerapan Algoritma Random Forest Untuk Prediksi Penjualan Dan Sistem Persediaan Produk," *Media Online*, vol. 5, no. 1, p. 20, 2024, doi: 10.30865/resolusi.v5i1.2149.
- [2] A. Ridwan, U. Muzakir, and S. Nurhidayati, "Optimizing E-commerce Inventory to prevent Stock Outs using the Random Forest Algorithm Approach," *Int. J. Softw. Eng. Comput. Sci.*, vol. 4, no. 1, pp. 107–120, Apr. 2024, doi: 10.35870/ijsecs.v4i1.2326.
- [3] N. Muntaja and S. Sriani, "Penerapan Algoritma Logika Fuzzy Mamdani Untuk Optimalisasi Stok Dari Berbagai Jenis Spareparts Handphone," *J. Comput. Syst. Informatics*, vol. 5, no. 4, pp. 1023–1032, 2024, doi: 10.47065/josyc.v5i4.5836.
- [4] M. H. Hamirsa and R. Rumita, "Usulan Perencanaan Peramalan (Forecasting) dan Safety Stock Persediaan Spare Part Busi Champion Type RA7YC-2 (EV-01/EW-01/2) Menggunakan metode Time Series Pada PT Triangle Motorindo Semarang," *Ind. Eng. Online J.*, vol. 11, no. 1, pp. 1–10, 2022, [Online]. Available: <https://ejournal3.undip.ac.id/index.php/ieoj/article/view/34373>.
- [5] F. A. Tuli, M. A. Mohammed, and D. K. Sachani, "Supply Chain Optimization: Machine Learning Applications in Inventory Management for E-Commerce," *Glob. Discl. Econ. Bus.*, vol. 13, no. 1, pp. 1–20, 2024, doi: 10.18034/gdeb.v13i1.758.
- [6] I. L. Kharisma, D. A. Septiani, A. Fergina, and K. Kamdan, "Penerapan Algoritma Decision Tree untuk Ulasan Aplikasi Vidio di Google Play," *J. Nas. Teknol. dan Sist. Inf.*, vol. 9, no. 2, pp. 218–226, 2023, doi: 10.25077/teknosi.v9i2.2023.218-226.
- [7] S. R. Febrian, A. A. Sunarto, and A. Pambudi, "PREDIKSI PENJUALAN SUKU CADANG MOTOR DENGAN PENERAPAN RANDOM FOREST DI PT TERUS JAYA SENTOSA MOTOR," 2024.
- [8] A. B. Trisnawan, P. Studi, S. Informasi, F. Sains, and M. Asia, "Integrasi Big Data dan Sistem Informatika Manufaktur dalam Prediksi Permintaan Produksi Integration of Big Data and Manufacturing Information Systems in Production Prediction," vol. 03, 2025.
- [9] A. T. Zy, Amali, A. M. Rifa'i, A. Z. Kamalia, and A. A. Sulaeman, "Detecting DDoS Attacks Through Decision Tree Analysis: An EDA Approach with the CIC DDoS 2019 Dataset," *2024 8th Int. Conf. Inf. Technol. Inf. Syst. Electr. Eng.*, pp. 202–207, 2024, doi: 10.1109/ICITISEE63424.2024.10730435.
- [10] R. Abraham *et al.*, "Forecasting a Stock Trend Using Genetic Algorithm and Random Forest. Journal of Risk and Financial Management," *J. Risk Financ. Manag.*, vol. 15, no. 5, 2022.
- [11] N. Sari, "Issn : 2809-7491," *J. Bisnis, Logistik dan Supply Chain*, vol. 2, no. 2, pp. 85–91, 2022.
- [12] M. Zakaria, S. Meutia, and A. Melinda Pane, "Perencanaan Produksi Dan Pengendalian Persediaan Bahan Baku Di Pt. Jakarana Tama Medan," *Ind. Eng. J.*, vol. 9, no. 2, 2020, doi: 10.53912/iejm.v9i2.574.

- [13] D. Arifuddin, “Comparison of the Performance of Multiple Linear Regression Algorithms and Multi Layer Perceptron Neural Networks in Predicting Drug Sales Perbandingan Performansi Algoritma Multiple Linear Regression dan Multi Layer Perceptron Neural Network dalam Mempred,” vol. 5, no. April, pp. 722–737, 2025.
- [14] F. Hamidy, “Optimalisasi Sistem Manajemen Persediaan untuk Pengendalian Stok yang Efisien Menggunakan Metode FIFO,” pp. 171–180, 2024.
- [15] A. Avinash, A. Widjaja, and O. Karnalim, “Analisis Perbandingan Algoritma Machine Learning untuk Forecasting Persediaan Produk Barang Pokok,” *J. Tek. Inform. dan Sist. Inf.*, vol. 10, no. 2, pp. 361–378, 2024, doi: 10.28932/jutisi.v10i2.9357.
- [16] A. R. Alamsyah, J. Jessen, J. Kharisma, R. Anggraini, and S. S. Muliati, “Analisa Penerapan Sistem Manajemen Operasional Pada PT. Unilever Tbk.,” *J. Ekon. Akutansi dan Manaj. Nusant.*, vol. 1, no. 2, pp. 72–79, 2023, doi: 10.55338/jeama.v1i2.20.
- [17] F. Fitriyadi and R. T. Saputra, “Visualisasi Tren Historis Bitcoin Menggunakan Python,” *J. Smart Syst.*, vol. 4, no. 2, pp. 12–20, 2025.
- [18] A. I. Zalukhu, M. Iqbal, and D. Nasution, “Analisis Data Mining Dalam Pengelolaan Persediaan Stok Dengan Algoritma Random Forest Dan Apriori (Studi Kasus: Toko Ceria Babyshop),” *J. Sci. Soc. Res.*, vol. 8, no. 3, pp. 3396–3405, 2025, [Online]. Available: <https://jurnal.goretanpena.com/index.php/JSSR/article/view/3544>.