

## Prediksi *Safety Stock* Menggunakan Algoritma *Support Vector Regression* Dengan Optimasi *Hyperparameter*

Rianindya Chandra Hardika<sup>1\*</sup>, Fiky Anggara<sup>2</sup>, Mohammad Ahmad Maidanul Abrori<sup>3</sup>

Sistem Informasi, Sekolah Tinggi Teknologi Bontang, Bontang, Indonesia

\*e-mail *Corresponding Author*: hardikachandra@gmail.com

### Abstract

*Inventory management plays an essential role in ensuring smooth supply chain operations and preventing both stockouts and overstocking. One effective approach to determining the appropriate level of safety stock is the use of machine learning-based prediction methods. This study aims to predict safety stock values using the Support Vector Regression (SVR) method with hyperparameter optimization through GridSearch. The dataset used is a public bike rental dataset, which includes variables such as time, weather, season, and holidays. The research stages include data preprocessing, Exploratory Data Analysis (EDA), implementation of the SVR model, and model performance evaluation using MAE, MAPE, and  $R^2$  metrics. The performance of the Support Vector Regression (SVR) algorithm with a Radial Basis Function (RBF) kernel and optimal parameters ( $C = 10$ ,  $\gamma = 0.001$ , and  $e = 0.01$ ) demonstrates strong predictive accuracy. Based on the Mean Absolute Percentage Error (MAPE), the model achieved prediction accuracies of 83.8%, 80.1%, and 81.7% on the training, validation, and testing datasets, respectively, indicating its effectiveness in modeling non-linear data. This model is capable of generating more precise safety stock predictions, thereby supporting decision-making in inventory planning and reducing the risk of stock shortages.*

**Keywords:** *Safety Stock; Support Vector Regression; Demand Prediction; Hyperparameter Tuning*

### Abstrak

Manajemen persediaan memiliki peran penting dalam menjaga kelancaran rantai pasok dan menghindari kekurangan maupun kelebihan stok. Salah satu pendekatan yang efektif untuk menentukan jumlah *safety stock* adalah dengan memanfaatkan metode prediksi berbasis *machine learning*. Penelitian ini bertujuan untuk memprediksi nilai *safety stock* menggunakan metode *Support Vector Regression* (SVR) dengan optimasi *hyperparameter tuning* melalui teknik *GridSearch*. Data yang digunakan merupakan dataset publik penyewaan sepeda yang mencakup variabel waktu, cuaca, musim, dan hari libur. Tahapan penelitian meliputi *preprocessing* data, analisis *Exploratory Data Analysis* (EDA), penerapan model SVR, serta evaluasi kinerja model menggunakan metrik MAE, MAPE, dan  $R^2$ . Hasil penelitian menunjukkan bahwa model SVR dengan kernel RBF dan parameter optimal ( $C=10$ ,  $\gamma=0.001$ ,  $e=0.01$ ) menghasilkan tingkat akurasi prediksi yang baik. Berdasarkan nilai MAPE, model mencapai akurasi 83,8% pada data latih, 80,1% pada data validasi, dan 81,7% pada data uji, yang mengindikasikan kemampuan SVR dalam memodelkan data non-linear. Model ini mampu memprediksi kebutuhan *safety stock* lebih baik, sehingga dapat membantu pengambilan keputusan dalam perencanaan persediaan dan mengurangi risiko kekurangan stok.

**Kata kunci:** *Safety Stock; Support Vector Regression; Prediksi Permintaan; Hyperparameter Tuning*

### 1. Pendahuluan

Manajemen persediaan, atau yang dikenal juga sebagai *inventory management*, merupakan elemen vital dalam operasional bisnis. Aktivitas ini mencakup berbagai fungsi penting seperti perencanaan, pengendalian, serta pengawasan terhadap stok barang yang dimiliki perusahaan. Dalam tahap perencanaan persediaan, faktor-faktor seperti permintaan pelanggan, hasil analisis data, tren pasar, serta aspek lain yang mempengaruhi permintaan menjadi perhatian utama. Manajemen inventori berperan penting dalam memastikan ketersediaan barang

yang tepat sekaligus menekan biaya akibat kelebihan atau kekurangan stok. Penerapan strategi *safety stock* yang efektif dapat membantu perusahaan memenuhi permintaan pasar, menjaga tingkat layanan pelanggan, dan mengoptimalkan penggunaan modal [1]. Beberapa faktor yang mempengaruhi pengelolaan persediaan antara lain fluktuasi permintaan pelanggan [1] [2], waktu tunggu (*lead time*) [3][4],[5][6],[7][8], perubahan harga [3][1][8], risiko pasokan [3][9][10][7] [11][12], serta biaya penyimpanan [3][1][12] [8]. Dalam konteks ini, *safety stock* memiliki peran strategis dalam menjaga kelancaran aliran barang serta memastikan ketersediaan bahan baku di setiap tahap produksi dan distribusi.

Dalam Praktiknya, Perusahaan sering menghadapi berbagai tantangan dalam kondisi permintaan yang tidak menentu, mempertahankan stok pengaman menjadi krusial untuk mencegah kekosongan barang (*stockout*). Estimasi yang tidak tepat terhadap jumlah *safety stock* dapat menimbulkan fenomena *bullwhip effect* (BWE) [9]. Beberapa penelitian sebelumnya menunjukkan bahwa untuk meminimalkan *bullwhip effect*, dapat digunakan pendekatan seperti jaringan syaraf buatan (*neural network*), model matematika, serta *data mining* [12]. Pendekatan-pendekatan tersebut juga dapat dimanfaatkan untuk memprediksi kebutuhan persediaan di masa mendatang namun hasilnya masih menunjukkan akurasi yang bervariasi. Hal ini mengindikasikan adanya permasalahan penentuan metode prediksi yang paling efektif untuk mendukung pengambilan keputusan persediaan.

Untuk mengatasi berbagai permasalahan dalam penentuan persediaan, diperlukan pendekatan prediksi permintaan yang mampu menangkap pola data yang kompleks dan dinamis [13]. *Machine learning* menjadi solusi yang relevan karena memiliki kemampuan memodelkan hubungan *non-linier*, memanfaatkan berbagai variabel eksternal seperti tren, musim, serta faktor cuaca, dan memberikan akurasi yang lebih baik dibandingkan metode konvensional [14]. Penelitian ini menggunakan *support vector regression* (SVR) sebagai model prediksi permintaan yang mampu menangani pola data yang bersifat *non-linear* dan fluktuatif [2]. Hasil prediksi dari model SVR tersebut kemudian diintegrasikan ke dalam perhitungan *safety stock*. Sebagai langkah antisipatif terhadap ketidakpastian hasil peramalan, perusahaan perlu menetapkan *safety stock* sebagai cadangan untuk menjaga kontinuitas pemenuhan permintaan. Sehingga jumlah persediaan pengaman dapat disesuaikan secara dinamis mengikuti perubahan kebutuhan aktual [7].

Tujuan penelitian ini adalah menerapkan algoritma SVR untuk prediksi permintaan dalam mendukung pengambilan keputusan pada manajemen persediaan. Hasil prediksi dari model SVR diintegrasikan ke dalam penentuan *safety stock*, sehingga dapat membantu meningkatkan efisiensi operasional, mengurangi risiko *stockout*, serta mengoptimalkan penggunaan modal melalui penetapan tingkat persediaan.

## 2. Tinjauan Pustaka

Dalam perkembangannya topik penelitian terkait dengan manajemen persediaan telah banyak dilakukan dengan berbagai metode. Penelitian sebelumnya telah dilakukan dengan hasil akurasi prediksi pada *lead time* kurang memuaskan dengan metode linear regresi, Ridge Regresi, Regresi Lasso, Regresi Pohon Keputusan, *random Forest* regresi, dan *k-Nearest Neighbors* dengan rata-rata nilai MAPE 38% [15]. Bing et al [8] melakukan prediksi perubahan persediaan masa depan pemasok eksternal dengan model *Backpropagation Neural Network* (BPNN) untuk menentukan *safety inventory* dibawah ketidakpastian dan mengatasi cacat jaringan BPPN menggunakan *Genetic Algorithm* (GA) untuk menyempurnakannya. Nilai MAPE menggunakan metode yang diusulkan adalah 3.63% lebih baik dibandingkan dengan metode tunggal. Data yang digunakan adalah Distributor *safety stock* Mewakili minimum jumlah persediaan yang distributor dapat menyimpannya berjalan normal, *retailer safety stock* merupakan total aman persediaan semua pengecer, *manufaktur inventory cost* mewakili biaya penyimpanan, kehilangan, pemeliharaan, *lead time* mewakili waktu dari pemesanan hingga pengiriman, *unit price* merupakan final jual harga satuan dari produk.

Bruno et al [16] mengkomparasi metode *machine learning* dan model jaringan syaraf untuk menentukan perkiraan pembelian produk untuk distributor farmasi. Data yang digunakan adalah riwayat penjualan inventaris lebih dari 10 ribu produk aktif selama 15 tahun terakhir. metode tersebut adalah *Seasonal Autoregressive Integrated Moving Average* (SARIMA), Prophet dari Facebook, *linear regression*, *Random Forest*, XGBoost and *Long Short-Term Memory* (LSTM). Hasil menunjukkan bahwa model SARIMA yang dirancang mengungguli model *machine learning* lainnya, dan model prediksi mingguan bekerja lebih baik daripada prediksi harian.

Peramalan model SARIMA pembelian produk harian dengan MAE rata-rata 9,38 untuk 300 produk teratas.

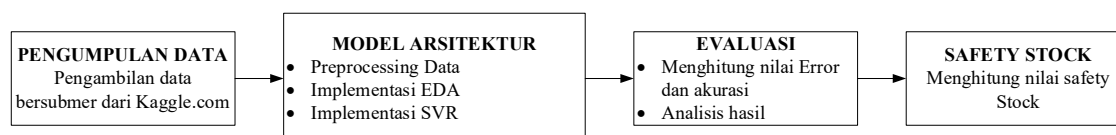
Penelitian [7] dengan model *Martingale of Forecast Evolution* (MMFE) mengembangkan pendekatan untuk menentukan tingkat persediaan pengaman berfokus pada lokasi penyimpanan, dari hasil lebih banyak ditempatkan di tiap hilir daripada hulu. Dua parameter yang menentukan alokasi *safety stock* adalah *lead time* dan level ketidakpastian permintaan. Semua stok pengaman harus ditempatkan pada tahap paling hilir di rantai pasokan. Ketika permintaan ketidakpastian rendah rata-rata sekitar 10% dari keamanan stok ditempatkan di hulu dalam rantai pasokan. Ketika tingkat ketidakpastian permintaan meningkat, ini menurun menjadi rata-rata sekitar 5%.

De Olivera [17] melakukan penelitian untuk mengurangi kesenjangan peramalan *lead time* untuk pembelian produk farmasi, perkiraan akurat dari *leadtime* tersebut dapat mengoptimalkan dan perencanaan dalam rantai pasok farmasi secara keseluruhan. Metode yang digunakan adalah *machine learning* dengan mengkomparasi 5 algoritma *k-nearest neighbors*, *support vector machines*, *random forests*, *linear regression* and *multilayer perceptrons*. Pendekatan yang diusulkan dalam penelitian ini dapat dibagi menjadi metode konvensional dan metode cerdas. Pertama tidak menggunakan kecerdasan buatan dan yang terakhir mengeksploitasi penambahan data dan pembelajaran mesin. Rata-rata SVM mengungguli metode lainnya dengan nilai MAPE 1.89%. Usulan pada penelitian ini adalah pendekatan yang lebih baik dapat memanfaatkan simulasi sistem *non-linear* untuk menyelidiki bagaimana kesalahan mesin dan prosedur dapat mempengaruhi *lead time*.

Dari berbagai penelitian yang telah dilakukan dalam upaya pengamanan persediaan, terlihat bahwa performa metode prediksi permintaan dan *lead time* masih menunjukkan variasi yang cukup besar. Namun, sebagian besar penelitian tersebut belum sepenuhnya menangkap dinamika data yang bersifat *non-linear* serta hanya berfokus pada peningkatan akurasi prediksi tanpa mengintegrasikannya secara langsung ke dalam kebijakan pengendalian persediaan. Dalam penelitian ini, dilakukan integrasi hasil prediksi menggunakan *Support Vector Regression* (SVR) ke dalam perhitungan *safety stock* untuk meningkatkan ketepatan estimasi persediaan pengaman sehingga ketersediaan barang tetap terjaga tanpa menimbulkan kelebihan stok. Pendekatan yang diusulkan memberikan keunggulan berupa prediksi permintaan yang lebih adaptif terhadap pola data *non-linear* serta mampu memberikan dampak langsung pada peningkatan akurasi penentuan jumlah *safety stock*.

### 3. Metodologi

Jalannya penelitian ini terdiri dari 4 Tahap yaitu mengumpulkan data, pembuatan arsitektur, evaluasi, dan nilai *Safety Stock*. Pengumpulan data dilakukan dengan mengambil data dari sumber dataset publik yaitu *kaggle.com*. dataset tersebut adalah data sewa sepeda dengan format CSV. Tahap kedua dari penelitian ini adalah membuat arsitektur, dalam tahap kedua yaitu proses *preprocessing* dilakukan pengolahan dan pembersihan data penyewaan sepeda, selanjutnya data diproses menggunakan *Exploratory Data Analysis* (EDA) tujuannya untuk memahami karakteristik, struktur, dan pola dalam dataset sebelum melakukan analisis lebih lanjut atau membangun model SVR. Proses data dengan model SVR untuk data *non-linear* dan ditingkatkan hasil prediksi dengan *hyperparameter tuning* yaitu *gridsearch*. Tahap ketiga dari penelitian ini adalah evaluasi, evaluasi akan dilakukan untuk menganalisis hasil nilai akurasi dari model yang dirancang. Terakhir menentukan nilai *safety stock* pada permintaan dimasa depan dengan menggabungkan nilai prediksi SVR dengan menambahkan *leadtime* waktu sewa sepeda sehingga suatu barang tersedia dan siap digunakan.



**Gambar 1.** Tahap Implementasi *Safety Stock*

Penelitian menggunakan SVR dalam upaya meningkatkan performa prediksi permintaan untuk menentukan nilai *safety stock*. Beberapa tahap penelitian dibutuhkan dalam mencapai hasil prediksi yang optimal dan rendah *error* seperti *preprocessing*, identifikasi variabel dengan EDA, implementasi SVR, evaluasi sistem dan nilai *safety stock* dari prediksi SVR.

### 3.1. Persiapan Data

Data yang diperoleh menggunakan dataset yang bersumber dari kaggle.com. data yang diambil merupakan data jumlah permintaan sewa sepeda berdasarkan waktu yang dipengaruhi oleh variabel cuaca, musim, liburan dan waktu buka penyewaan. data yang diperoleh diantara 01 Desember 2017 sampai dengan 30 November 2018. Data tersebut dalam bentuk file *Comma Separated Value* (CSV).

### 3.2. Preprocessing data

pada tahap ini dilakukan pengolahan data agar lebih mudah digunakan untuk proses selanjutnya. Pembersihan data dilakukan mencakup penyesuaian dan penghapusan data. Data yang diperoleh menghitung permintaan berdasarkan waktu dan tanggal sehingga digabungkan menjadi satu permintaan setiap harinya. Pada variabel *functioning day* yang bernilai "No" maka tidak ada sepeda yang disewa sehingga dapat dihapus. Melakukan perubahan data *type* data setiap variabel seperti mengubah datetime dengan format *Year-Month-Day*. Selanjutnya mengisi data yang kosong menjadi 0.

### 3.3. Identifikasi EDA

Identifikasi dalam Konteks EDA mengacu pada langkah awal dalam analisis data dimana tujuannya untuk mengenali dan memahami karakteristik, struktur dan pola dalam dataset *SeoulBikeRent.CSV*. Selanjutnya melakukan perhitungan korelasi antara dua variabel numerik pada sewa sepeda dengan pengaruh faktor cuaca yaitu *rented Bike Count*, *Temperature*, *Humidity*, *wind speed*, *visibility*, *Dew Point Temperature*, *solar Radiation*, *Rainfall*, dan *Snowfall*.

### 3.4. Support Vector Regression (SVR)

*Support vector regression* (SVR) adalah algoritma dalam bidang pembelajaran mesin yang dirancang untuk memetakan hubungan antara variabel bebas dan terikat pada permasalahan regresi, dengan kemampuan tinggi dalam menangani pola data yang kompleks serta bersifat *non-linier* [2].

Fungsi regresi SVR pada kasus linear dengan persamaan berikut:

$$f(x_i) = \sum_{i=1}^n (a_i - a_i^*) (x_i \cdot x) + b \quad (1)$$

Fungsi regresi SVR untuk non linear dinotasikan dengan persamaan berikut:

$$(x_i) = \sum_{i=1}^n (a_i - a_i^*) \Phi(x_i) \Phi(x_j) + b \quad (2)$$

Setelah identifikasi EDA adalah SVR. Dalam model SVR variabel kategorik dilakukan *encoding* menggunakan *One-hot Encoding* berguna untuk mengubah variabel yang dapat digunakan untuk model SVR. Langkah yang dilakukan adalah identifikasi variabel kategori. Variabel pada penelitian yang dilakukan adalah *seasons* (*winter*, *spring*, *autumn*, *Summer*) dan *holiday* (*No Holiday*, *Holiday*). Selanjutnya melakukan *one-Hot Encoding* untuk setiap variabel kategori. Proses ini melibatkan transformasi kolom baru untuk setiap kategori unik dalam variabel. kolom akan berisi nilai *biner* (0 Atau 1) yang menunjukkan apakah kategori tersebut ada atau tidak dan variabel numerik distandarisasi menggunakan *MinMax Scaler* digunakan pada data numerik seperti keadaan cuaca (*solar radiation*, *temperature*), tujuan menggunakan *MinMax Scaler* mengubah data numerik ke dalam rentang tertentu sebelum digunakan dalam model SVR. Harapannya dapat meningkatkan kinerja model SVR terutama jika fitur memiliki rentang nilai yang sangat berbeda. *MinMax Scaler* bekerja dengan mengubah setiap fitur numerik menjadi rentang tertentu. Dengan perhitungan tersebut dengan *MinMax Scaled* didapatkan nilai yang sudah dinormalisasi. Sehingga data fitur *wind speed* dapat sebagai *input* untuk model SVR.

### 3.5. Hyperparameter Tuning

*Hyperparameter tuning* dilakukan untuk mencari kombinasi *hyperparameter* yang optimal pada model SVR. *Hyperparameter tuning* yang digunakan adalah *gridsearch* yang merupakan metode sistematis dengan mencoba kombinasi C, *Kernel* dan *degree*, *epsilon*, *gamma*. Kombinasi C dengan rentang 0,001, 1, 10, 100, 1000. Kombinasi *kernels* yang digunakan *polynomial*, *Radial Basis Function* (RBF), dan *Sigmoid*. Kombinasi *degree* dengan 3, 4, 6, 8, 10.

### 3.6. Prediksi Safety Stock

*Safety stock* merupakan sejumlah persediaan tambahan yang disiapkan untuk mengantisipasi adanya ketidaksesuaian antara hasil peramalan dan permintaan aktual, termasuk perbedaan antara *delivery time* yang diharapkan dan aktualnya, serta faktor-faktor tidak terduga lainnya [7]. Besaran *safety stock* yang dibutuhkan untuk memenuhi tingkat permintaan tertentu dapat ditentukan melalui simulasi atau metode statistik. Rumus yang digunakan untuk menghitung nilai *safety stock* adalah [5] :

$$SS = (M_D \times M_{lt}) - (A_D \times A_{lt}) \quad (3)$$

Keterangan:

*SS* = Jumlah *safety stock* yang dibutuhkan

*MD* = Maksimal Penggunaan perhari

*Mlt* = Maksimal *lead time*

*AD* = Rata-rata penggunaan perhari

*Alt* = Rata-rata *lead time*

### 3.7. Metode Evaluasi

Metode penilaian digunakan dalam penelitian ini untuk mengukur keakuratan metode peramalan dan mengevaluasi kinerja algoritma SVR. Analisis prediksi digunakan untuk memperkirakan nilai di masa mendatang berdasarkan pola yang diperoleh dari data historis. Keberhasilan suatu model peramalan dapat dinilai melalui tingkat kesalahan atau *forecast error* yang menunjukkan perbedaan antara nilai aktual dan hasil prediksi [20]. Dalam penelitian ini, menggunakan metrik evaluasi untuk menilai kinerja algoritma SVR dalam menghasilkan prediksi yang akurat.

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (4)$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (5)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (6)$$

## 4. Hasil dan Pembahasan

Hasil penelitian ini menyajikan hasil dari *preprocessing* data, identifikasi EDA, Implementasi prediksi data non-linear model SVR Nilai prediksi tersebut menjadi variabel nilai prediksi *safety stock* pada sewa sepeda.

### 4.1. Persiapan dan Preprocessing Data

Penelitian ini dimulai dengan mencari data yang memiliki pola data non-linear yang berbasis permintaan. Data yang diperoleh pada dataset publik yaitu adalah *kaggle.com* dengan rincian dataset pada table 1.

**Tabel 1.** Dataset *seoulBikedrent.csv*

Date	Temp (°C)	Solar Radiation (MJ/m2)	Rainfall (mm)	Demand	Seasons	Holiday
2017-12-01	-2.45	0.24875	0	9539	Winter	No Holiday
2017-12-02	1.325	0.26375	0	8523	Winter	No Holiday
2017-12-04	-0.304	0.282917	0.004	8729	Winter	No Holiday
....	....	....	...	...	...	...
2018-11-29	3.304	-9.8083	0,117	16423	Autumn	No Holiday
2018-11-30	2.7625	-8.37083	0.425	16297	Autumn	No Holiday

Data yang diperoleh adalah *seoulBikedrent.csv* dengan rentang waktu 01 Desember 2017 sampai dengan 30 November 2018 dengan 8760 Baris data dengan fungsi *data.describe().T* untuk menghasilkan statistik deskriptif ringkas dari data setiap kolom dengan tujuan memudahkan dalam memahami data yang akan digunakan dalam proses prediksi.

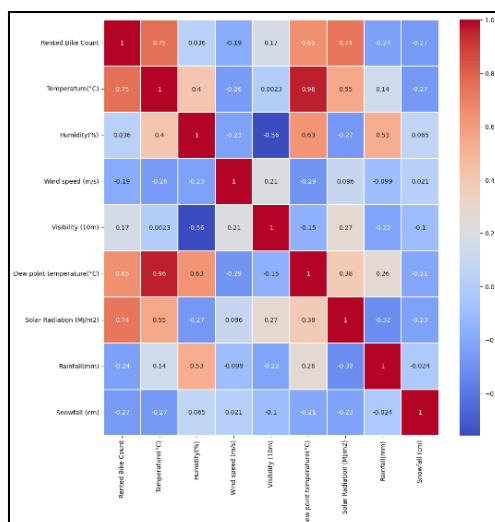
**Tabel 2.** Ringkasan statistik pada dataset faktor cuaca

Fitur	Count	Mean	Std	Min	25%	50%	75%	Max
Hour	8760	11.500	6.9225	0.0	5.75	11.50	17.25	23.00
Count	8760	704.602	644.99	0.0	191.00	504.50	1065	3556
Temperature(°C)	8760	12.882	11.944	-17.8	3.50	13.70	22.50	39.40
Humidity(%)	8760	58.2262	20.362	0.0	42.00	57.00	74.00	98.00
Wind speed (m/s)	8760	1.724909	1.036	0.0	0.90	1.50	2.30	7.40
Visibility (10m)	8760	1436.825	608.29	27.0	940.00	1698.00	2000	2000
Dew Point Temperature (C)	8760	4.07381	13.060	-30.6	-4.70	1698.00	14	27.20
Solar Radiation (MJ/m2)	8760	0.5691	0.8687	0.0	0.00	0.01	0.93	35.003
Rainfall (mm)	8760	0.1486	1.128	0.0	0.00	0.00	0.00	35.00
Snowfall (cm)	8760	0.0750	0.436	0.0	0.00	0.00	0.00	8.80

Tabel 2 menunjukkan ringkasan statistik dari berbagai faktor cuaca dalam dataset yang berjumlah 8.760 data. Ringkasan ini memberikan gambaran awal mengenai karakteristik dan variasi kondisi cuaca pada dataset. Selain itu melakukan perhitungan jumlah sewa sepeda pada *functioning day*. Didapatkan nilai nol jika *functioning Day No* dan *functioning day yes* dengan jumlah sepeda yang disewakan 6172314. Maka *functioning day "no"* dapat dihapus baris data karena tidak berpengaruh pada data. Melakukan perubahan data *type* data setiap variabel mengubah *datetime* dengan format *Year-Month-Day*. Selanjutnya mengisi data yang kosong menjadi 0.

#### 4.2. Identifikasi EDA

Dalam proses *Exploratory Data Analysis* (EDA) dapat melihat distribusi, korelasi dan karakteristik penting dalam dataset. Data yang akan dilihat adalah *Rented Bike Count*, *temperature*, *humidity*, *wind speed*, *visibility*, *dew point temperature*, *solar radiation*, *rainfall*, dan *snowfall*. Dari gambar 3 tiga variabel dengan nilai korelasi yang tinggi dengan variabel *rented bike count* adalah *solar radiation*, *temperature*, *dew point temperature*. Akan tetapi *temperature* dan *dew point temperature* memiliki korelasi tinggi dapat mengakibatkan multikolinearitas. multikolinearitas dapat menjadi masalah dalam analisis regresi karena dapat mengganggu interpretasi hasil model dan membuat koefisien regresi menjadi tidak stabil sehingga variabel *dew point temperature* akan dihapus.

**Gambar 2.** Heatmap EDA faktor cuaca

Pengujian pada tabel kategori *holiday* dilakukan dengan uji-T, hasil nilai *P-Value*  $Ttest\_indResult(statistic=-2.0444149000230416, pvalue=0.04165827462260651)$  yang lebih kecil dari 0.05 sehingga terdapat perbedaan yang signifikan antara jumlah sepeda yang disewa

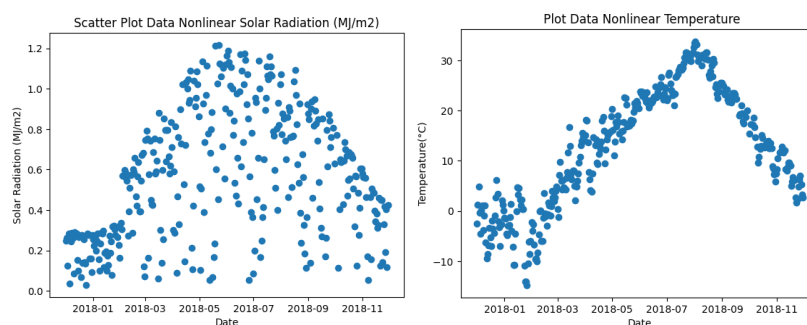
pada hari libur dan bukan hari libur. Pada data *seasons* dilakukan uji kruskal *wallisKruskalResult*(statistic=188.39660365763575,pvalue=1.3550678519480538e-40), nilai P-Value kecil dari 0.05 dengan kata lain ada perbedaan signifikan data musim yang mempengaruhi jumlah sepeda yang dirental.

Berdasarkan hasil EDA terdapat dua variabel numerik yang memiliki korelasi tinggi terhadap *rented bike count* yaitu *solar radiation* dan *temperature*. Sedangkan data kategori variabel *seasons* dan *holiday* memiliki pengaruh terhadap jumlah sepeda yang dirental. Dalam tahap implementasi model data akan dibagi menjadi dua yaitu data latih dan uji. Karena datanya *time series* maka tidak dapat dilakukan secara *random* membagi data. Oleh karena itu data akan menjadi data uji dari Desember 2017 – September 2018 dan data uji dari oktober 2018 – Desember 2018.

**Tabel 3.** Rentang dataset latih dan uji

Data	Rentang waktu	Jumlah data
Data latih	Desember 2017 – September 2018	298
Data Uji	Oktober 2018 – Desember 2018.	55

Untuk membuktikan bahwa data cuaca memiliki pola data *non linear* dengan menggunakan *scatter plot* data dari *library Matplotlib*. Dalam data linear perubahan dalam satu variabel cenderung berbanding lurus atau sebanding dengan perubahan variabel lainnya dan data *non-linear* perubahan dalam satu variabel hubungannya berubah-ubah [18]. Pada gambar 3 menunjukan variabel *Solar Radiation* dan *temperature* tidak garis lurus sehingga dapat dikatakan bahwa data cuaca adalah data *non linear*.



**Gambar 3.** Plot data *non-linear* pada data cuaca

#### 4.3. Implementasi model SVR

Dalam pemodelan SVR, variabel kategorik dilakukan *encoding* menggunakan *one hot encoding* dan variabel numerik distandarisasi menggunakan *minmax scaler*. Tabel 4 adalah sebagian hasil data dari proses *one hot encoding* dan *minmax scaler*.

**Tabel 4.** Hasil *one hot encoding* dan *minmax scaler*

<i>Solar_scaled</i>	<i>Temp_scaled</i>	<i>Season_Winter</i>	<i>Season_Summer</i>	<i>Holiday_No</i>	<i>Holiday_Yes</i>
-1.22	-1.14	1	0	1	0
0.00	-0.13	0	1	0	1
1.22	1.27	1	0	1	0

Berdasarkan hasil pada tahap identifikasi EDA variabel yang berpengaruh adalah *Solar radiation*, *temperature*, *seasons*, dan *holiday*. variabel tersebut akan di *transform* untuk mengubah skala dan menormalkan data sebelum digunakan di algoritma SVR dan variabel kategori menggunakan teknik transformasi *one hot encode*. Implementasi pada tabel 2 adalah hasil prediksi SVR dengan linear kernel atau *default* Kernel pada SVR. Peningkatan yang terlalu tinggi pada nilai MAPE antara data latih dan data uji mengindikasikan bahwa model mengalami *overfitting* atau tidak dapat menggeneralisasi dengan baik ke data yang belum dilihat sebelumnya. Kenaikan ini menjadi peringatan bahwa model tidak dapat diandalkan dalam membuat prediksi yang akurat pada data baru.

**Tabel 5.** Hasil SVR tidak menggunakan *hyperparameter GridSearch*

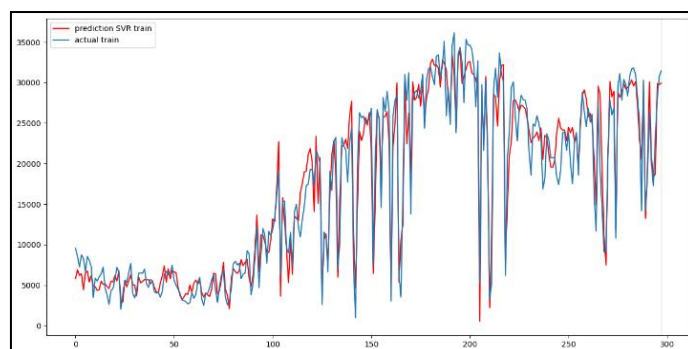
Data	MAE	MAPE	R2	RMSE
Data latih	1739.713	0.162	0.94	2520.177
Data Uji	6195.913	0.359	-0.554	6696.486

Dari tabel 5 diperoleh nilai MAPE data latih 16.2% dan data uji 35.9%. Nilai ini akan dibandingkan dengan penggunaan *hyperparameter tuning* dengan menggunakan *gridsearch* dalam mencari kombinasi terbaik dari parameter model. Parameter C (*Cost parameter*) [0.001, 0.01, 1, 10, 100, 1000], kernels ['poly', 'rbf', 'sigmoid'], degrees untuk *kernel Polynomial* [3, 4, 6, 8, 10], gamma = [0.001, 0.01, 0.1, 1, 10], dan *epsilon* = [0.01, 0.1, 0.2].

**Tabel 6.** Hasil Metode SVR pada data latih dengan *gridsearch*

Rented Bike Count	Solar Radiation (MJ/m2)	Temperature(°C)	Seasons	Holiday	Date	SVR Prediction
9539	0.248750	-2.454.167	Winter	No Holiday	01/12/2017	5798.0
8523	0.263750	1.325.000	Winter	No Holiday	02/12/2017	6884.0
...	...	...	...	...	...	...
26398	0.794167	21.008.333	Autumn	No Holiday	22/09/2018	30071.0
20060	0.615417	19.412.500	Autumn	Holiday	23/09/2018	20938.0

Hal ini menunjukkan bahwa *hyperparameter* dapat menurunkan nilai MAPE dengan model SVR. *Hyperparameter* terbaik adalah nilai C= 10, Kernel = RBF, dan gamma = 0.001, dan epsilon 0.01 dengan data validasi MAPE sebesar 16.2%. Pada gambar 5 adalah grafik data latih yang dibandingkan dengan data aktual. Nilai MAPE pada data latih adalah 16.2%.

**Gambar 4.** Prediksi model SVR pada data latih dengan *gridsearch*

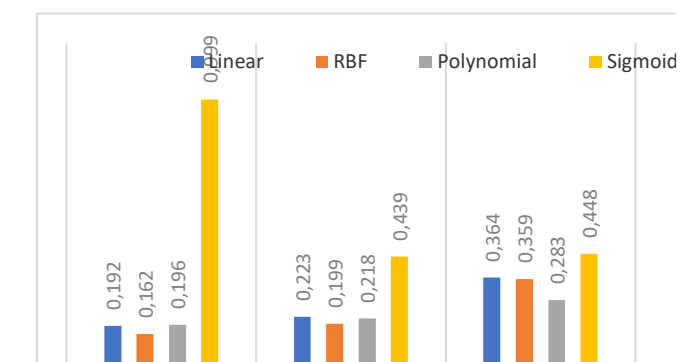
Pada grafik setelah dilakukan proses *hyperparameter tuning* menggunakan *gridSearch* terlihat bahwa garis prediksi mengikuti pola garis data aktual, menunjukkan bahwa model mampu mempelajari hubungan antar variabel dengan baik. Hasil prediksi dari data latih menggunakan *hyperparameter tuning* yaitu *gridSearch* menghasilkan nilai C= 10, Kernel = RBF, dan gamma = 0.001, dan epsilon 0.01.

**Tabel 7.** Hasil Metode SVR pada data uji dengan *gridsearch*

Rented Bike Count	Solar Radiation (MJ/m2)	Temp (°C)	Seasons	Holiday	Date	SVR prediction
19877	0.410000	7.395833	Autumn	No Holiday	2018-10-29	19132.1
20822	0.606250	5.891667	Autumn	No Holiday	2018-10-30	19568.2
22964	0.581250	8.345833	Autumn	No Holiday	2018-11-01	21814.7
24077	0.563333	10.504167	Autumn	No Holiday	2018-11-02	23620.5

Hasil *hyperparameter* adalah yang terbaik dibandingkan dengan kernel *default* SVR kernel linear dengan perbandingan 36.4% dan 35.9%. Gambar 6 tersebut menunjukkan

perbandingan nilai MAPE dari empat jenis kernel yang digunakan pada algoritma SVR, yaitu Linear, RBF, Polynomial, dan Sigmoid, pada tiga skenario pengujian data yaitu data latih, data validasi, dan data uji.



Gambar 5. perbandingan MAPE terbaik dari data validasi

Secara umum terlihat bahwa kernel RBF memberikan performa terbaik, khususnya pada data latih dengan nilai MAPE terendah yaitu 0,162, yang menunjukkan bahwa kernel ini mampu mempelajari pola permintaan yang bersifat non-linear dengan akurasi tinggi. Keunggulan kernel RBF masih konsisten ketika diuji pada data validasi dengan MAPE sebesar 0,199, mengindikasikan bahwa model tidak hanya baik dalam pelatihan tetapi juga memiliki kemampuan generalisasi yang lebih baik dibanding kernel lainnya. Sementara itu, kernel Linear dan Polynomial memiliki performa yang relatif mirip pada semua skenario pengujian, namun cenderung memiliki tingkat kesalahan prediksi lebih tinggi dibanding RBF.

#### 4.4. Evaluasi Model

Pada penelitian ini, model SVR digunakan untuk memprediksi permintaan sebagai dasar penentuan *safety stock*. Proses peningkatan performa model dilakukan melalui teknik *hyperparameter tuning* menggunakan *GridSearch*. Pada tabel 8 adalah hasil *hyperparameter tuning* dengan menggunakan *gridsearch* dari data validasi. Berdasarkan pencarian parameter terbaik didapatkan nilai dengan C 10.0, kernel RBF, *gamma* 0.001 dan *epsilon* 0.01. Nilai validasi MAPE terbaik dengan nilai 21%

Tabel 8. *Hyperparameter tuning* dengan *Grid Search*

C	kernel	Gamma	Epsilon	Degree	Validation MAE	Validation MAPE	Validation R2
1,0	linear	0,001	0,01	-	6610,84	0,278	-0,372
1,0	poly	0,001	0,01	3	5070,07	0,218	0,200
10,0	rbf	0,001	0,01	-	4880,34	0,199	0,230
0,001	sigmoid	0,001	0,2	-	8686,34	0,439	-1,265

Berdasarkan hasil *GridSearch*, diperoleh kombinasi *hyperparameter* terbaik berupa C = 10.0, kernel RBF, *gamma* = 0.001, dan *epsilon* = 0.01. Kombinasi parameter tersebut menghasilkan nilai MAPE pada data validasi sebesar 0.199 atau sekitar 19,9%. Jika nilai MAPE dikonversi ke dalam skala akurasi, model memiliki tingkat akurasi sebesar 80,1%. Hal ini mengindikasikan bahwa model yang telah melalui proses *tuning* dapat menghasilkan prediksi permintaan yang cukup baik pada data yang tidak termasuk dalam proses pelatihan.

Tabel 9. Perbandingan nilai SVR evaluasi *gridsearch*

Data	Model	MAE	MAPE	R2
Latih	SVR	1927.157	0.192	0.936
	SVR-Gridsearch	1739.713	0.162	0.94
Uji	SVR	5891.69	0.364	-0.372
	SVR-Gridsearch	2574.175	0.183	0.596

Evaluasi selanjutnya dilakukan dengan membandingkan performa SVR sebelum dan sesudah *tuning* pada data latih dan data uji. Pada data latih, SVR dengan *GridSearch* menunjukkan peningkatan kinerja, di mana nilai MAPE turun dari 0.192 menjadi 0.162. Konversi nilai MAPE tersebut menunjukkan bahwa akurasi prediksi meningkat dari 80,8% menjadi 83,8%. Hasil ini menegaskan bahwa *tuning* parameter berhasil meningkatkan ketepatan model dalam mengenali pola data historis.

Pada data uji, performa model juga menunjukkan peningkatan yang signifikan setelah dilakukan *tuning*. Model SVR memiliki nilai MAPE sebesar 0,364 dengan akurasi 63,6% dan nilai  $R^2$  yang negatif (-0,372), mengindikasikan bahwa model tersebut tidak mampu memprediksi data baru dengan baik. Setelah *tuning*, nilai MAPE menurun menjadi 0,183 atau setara dengan akurasi 81,7%, dan nilai  $R^2$  meningkat menjadi 0,596.

#### 4.5. Safety Stock

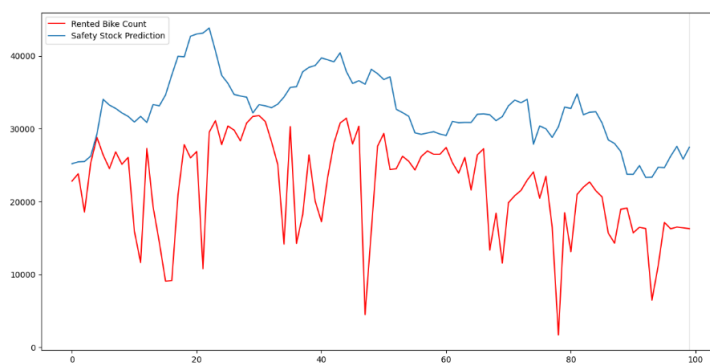
Dalam penelitian ini menghasilkan prediksi banyaknya permintaan dengan model SVR sehingga bisa menggunakan hasil prediksi yang dihasilkan SVR untuk melakukan prediksi *safety stock*. Selain itu, penelitian yang dilakukan dapat memprediksi banyaknya permintaan sehingga dapat memprediksi *safety stock* untuk dimasa depan.

**Tabel 10.** Hasil Prediksi *Safety Stock*

Date	Rented Bike Count	SVR prediction	Safety stock predict
2018-10-19	27437	26829,17079	29075,42857
2018-10-20	25355	27047,36533	31011,28571
.....	.....	.....	.....
2018-11-25	11212	17604,79795	24725,14286
2018-11-26	17162	20488,84804	24671,14286

Gambar 6 menampilkan grafik perbandingan antara jumlah permintaan aktual dengan hasil perhitungan *safety stock* berbasis prediksi permintaan. Grafik tersebut menunjukkan bahwa nilai *safety stock* yang dihasilkan model mengikuti pola fluktuasi permintaan yang terjadi. Ketika permintaan meningkat secara signifikan, nilai *safety stock* juga meningkat sehingga risiko kekurangan persediaan (*stockout*) dapat diminimalkan.

Di sisi lain, terlihat bahwa terdapat periode ketika *safety stock* turun cukup rendah dibanding nilai permintaan aktual. Kondisi ini menunjukkan bahwa *safety stock* disesuaikan dengan prediksi permintaan, bukan disesuaikan pada angka tinggi secara konstan.



**Gambar 6.** Grafik prediksi *safety stock*

Hasil prediksi permintaan yang dilakukan dengan model SVR menunjukan hasil yang memuaskan pada penentuan nilai *safety stock*. Penentuan nilai *safety stock* menunjukan dapat menangani permasalahan pada lonjakan permintaan yang tidak terduga terlihat pada gambar 7. Akan tetapi nilai *safety stock* yang dihasilkan selisih dengan data aktual cukup besar karena menggunakan *leadtime* dengan asumsi sepeda disewa rata-rata 1 hari dan perbaikan sepeda 1 hari sehingga total maksimal *lead time* adalah 2 hari. Perhitungan rata-rata harian menggunakan siklus 7 hari/mingguan mendapatkan nilai *safety stock* yang sangat baik dalam menyediakan stok pada *inventory* dibandingkan menggunakan perhitungan bulanan atau harian.

#### 4.6. Pembahasan

Berbagai penelitian prediksi *safety stock* menunjukkan bahwa metode konvensional dan model *machine learning* tanpa optimasi parameter memiliki keterbatasan dalam menangani fluktuasi permintaan harian yang bersifat *non-linear*, yang berdampak pada tingginya kesalahan prediksi persediaan. Penggunaan algoritma regresi berbasis *machine learning* mampu menurunkan tingkat kesalahan dibandingkan pendekatan statistik tradisional [16], namun akurasi masih bergantung pada pemilihan dan penyesuaian parameter model. Penelitian ini menunjukkan bahwa algoritma SVR dengan *hyperparameter tuning* mampu menurunkan kesalahan prediksi dengan nilai akurasi 83,8% pada data latih dan akurasi 81,7% pada data uji berdasarkan evaluasi MAPE. Temuan ini menunjukkan kontribusi pengembangan metode prediksi *safety stock* khususnya untuk permintaan harian yang memiliki pola *non-linear*. Namun demikian, keterbatasan jumlah data yang digunakan menyebabkan model belum sepenuhnya mampu mengenali pola permintaan secara optimal.

Penelitian ini juga menegaskan bahwa data permintaan yang bersifat *non-linear* dan dipengaruhi variabel cuaca dan musiman tidak tepat diprediksi dengan metode linier. Penelitian terdahulu menggunakan metode regresi linier, *Random Forest Regression*, dan *KNN Regression* menghasilkan MAPE yang dikategorikan rendah untuk kebutuhan operasional prediksi persediaan [19]. Penerapan SVR dengan kernel *non-linear* RBF dalam penelitian ini terbukti lebih adaptif dalam menangkap dinamika variabel eksternal, dan mampu menangani kompleksitas perilaku permintaan. Selain itu, penelitian ini menguatkan penelitian terdahulu [7] pentingnya akurasi prediksi untuk menentukan jumlah *safety stock* yang tepat. Dengan menghasilkan prediksi yang lebih akurat, nilai *safety stock* dari penelitian ini terbukti lebih responsif terhadap perubahan permintaan yang belum menjadi fokus di sebagian besar penelitian sebelumnya. Hasil tersebut menunjukkan bahwa keberhasilan model ini secara langsung berdampak pada pengurangan risiko *bullwhip effect* serta menekan potensi *stockout*.

#### 5. Simpulan

Hasil dari penelitian ini dapat memberikan solusi untuk penelitian dan pengembangan prediksi permintaan harian. Dengan menggunakan model SVR dapat memberikan hasil akurasi lebih baik. Oleh karena itu metode SVR dapat digunakan untuk mengatasi ketidakpastian permintaan. *Hyperparameter tuning* dengan *gridSearch* menghasilkan nilai prediksi yang lebih baik, yang dapat disimpulkan bahwa dataset bekerja lebih baik pada data *non-linear*. Hasil regresi menggunakan data linear tidak cukup baik dalam melakukan prediksi karena karakteristik data adalah non-linear.

Pada tahap pemodelan peningkatan performa model dilakukan melalui *hyperparameter tuning* menggunakan *GridSearch*, yang menghasilkan parameter terbaik  $C=10$ ,  $kernel=RBF$ ,  $gamma=0.001$ , dan  $epsilon=0.01$ . Model ini memberikan hasil yang lebih baik pada data latih dan uji dengan MAPE masing-masing 16,2%, 18,3%, meskipun performa pada data uji masih menunjukkan ketidakstabilan. Penelitian selanjutnya perlu mempertimbangkan metode optimasi parameter yang optimal atau eksplorasi teknik kombinasi untuk meningkatkan performa prediksi yang dapat meningkatkan akurasi prediksi yang lebih baik. Selain itu, Mengatasi masalah heteroskedastisitas pada deret waktu yang tidak konstan sepanjang rentang data perlu dilakukan karena Heteroskedastisitas dapat menjadi masalah dalam analisis statistik dan menyebabkan kesalahan dalam estimasi parameter model.

#### Daftar Referensi

- [1] E. Fatma and D. S. Pulungan, "Analisis Pengendalian Persediaan Menggunakan Metode Probabilistik dengan Kebijakan Backorder dan Lost sales," *J. Tek. Ind.*, vol. 19, no. 1, p. 38, 2018, doi: 10.22219/jtiumm.vol19.no1.40-51.
- [2] S. Jaipuria and S. S. Mahapatra, "A hybrid forecasting technique to deal with heteroskedastic demand in a supply chain," *Oper. Supply Chain Manag.*, vol. 14, no. 2, pp. 123–132, 2021, doi: 10.31387/oscm0450291.
- [3] D. A. Efrilianda, Mustafid, and R. R. Isnanto, "Inventory control systems with safety stock and reorder point approach," *2018 Int. Conf. Inf. Commun. Technol. ICOI ACT 2018*, vol. 2018-Janua, pp. 844–847, 2018, doi: 10.1109/ICOI ACT.2018.8350766.
- [4] C. J. Lee and S. C. Rim, "A Mathematical Safety Stock Model for DDMRP Inventory Replenishment," *Mathematical Problems in Engineering.*, vol. 2019, pp. 1-10, 2019, doi: 10.1155/2019/6496309.

- [5] T. Gong, J. Xu, W. Bian, and Z. Li, "Optimization of two-echelon supply chain metering device safety stock placement with uncertain demand," *14th Int. Conf. Serv. Syst. Serv. Manag. ICSSSM 2017 - Proc.*, pp. 1–6, 2017, doi: 10.1109/ICSSSM.2017.7996162.
- [6] G. Han and X. A. Sun, "A research on the storage strategy of university apartment spare parts based on the stochastic inventory model," *2016 Int. Conf. Logist. Informatics Serv. Sci. LISS 2016*, pp. 1–4, 2016, doi: 10.1109/LISS.2016.7854353.
- [7] Y. Boulaksil, "Safety stock placement in supply chains with demand forecast updates," *Oper. Res. Perspect.*, vol. 3, pp. 27–31, 2016, doi: 10.1016/j.orp.2016.07.001.
- [8] W. Bing and P. Chen, "Research on Supply Chain safety Inventory forecast based on GA-BP Neural Network," *RASSE 2021 - IEEE Int. Conf. Recent Adv. Syst. Sci. Eng. Proc.*, pp. 1–5, 2021, doi: 10.1109/RASSE53195.2021.9686846.
- [9] T. Inprasit and S. Tanachutiwat, "Reordering point determination using machine learning technique for inventory management," *ICEAST 2018 - 4th Int. Conf. Eng. Appl. Sci. Technol. Explor. Innov. Solut. Smart Soc.*, 2018, doi: 10.1109/ICEAST.2018.8434473.
- [10] Y. Cai, M. F. Zhang, and L. Huang, "Safety stock management based on lead time optimization," *2011 Int. Conf. Electr. Technol. Civ. Eng. ICETCE 2011 - Proc.*, pp. 6150–6153, 2011, doi: 10.1109/ICETCE.2011.5774201.
- [11] K. Jakkraphobyothin, S. Srita, and T. Chinda, "Factor Analysis of Inventory Management in Thai Construction Industry," *TIMES-iCON 2018 - 3rd Technol. Innov. Manag. Eng. Sci. Int. Conf.*, pp. 1–5, 2019, doi: 10.1109/TIMES-iCON.2018.8621817.
- [12] B. B. Barrios, A. A. Juan, J. Panadero, K. Altendorfer, A. J. Peirleitner, and A. Estrada-Moreno, "On the use of Simheuristics to Optimize Safety-Stock Levels in Material Requirements Planning with Random Demands," *Proc. - Winter Simul. Conf.*, vol. 2020-Decem, no. 2017, pp. 1539–1550, 2020, doi: 10.1109/WSC48552.2020.9383988.
- [13] P. R. Hakim and H. Prastawa, "Forecasting Demand & Usulan Safety Stock Pasir Silika Dengan Metode Time Series Pada Pt Solusi Bangun Indonesia Tbk. PABRIK ...," *Ind. Eng. Online J.*, vol. 11, no. 4, 2022, [Online]. Available: <https://prosiding.seminar-id.com/index.php/sainteks>
- [14] A. Jain, V. Karthikeyan, B. Sahana, S. Br, K. Sindhu, and S. Balaji, "Demand Forecasting for E-Commerce Platforms," *2020 IEEE Int. Conf. Innov. Technol. INOCON 2020*, pp. 9–12, 2020, doi: 10.1109/INOCON50539.2020.9298395.
- [15] S. Singh, "production system using various Machine Learning," *2019 9th Int. Conf. Cloud Comput. Data Sci. Eng.*, pp. 422–425, 2019.
- [16] B. K. Almentero, J. Li, and C. Besse, "Forecasting pharmacy purchases orders," *Proc. 2021 IEEE 24th Int. Conf. Inf. Fusion, FUSION 2021*, pp. 1–8, 2021, doi: 10.23919/fusion49465.2021.9627017.
- [17] M. B. de Oliveira, G. Zucchi, M. Lippi, D. F. Cordeiro, N. R. da Silva, and M. Iori, "Lead Time Forecasting with Machine Learning Techniques for a Pharmaceutical Supply Chain," *Int. Conf. Enterp. Inf. Syst. ICEIS - Proc.*, vol. 1, no. Iceis, pp. 634–641, 2021, doi: 10.5220/0010434406340641.
- [18] K. Goswami and A. B. Kandali, "Electricity Demand Prediction using Data Driven Forecasting Scheme: ARIMA and SARIMA for Real-Time Load Data of Assam," *2020 Int. Conf. Comput. Perform. Eval. ComPE 2020*, pp. 570–574, 2020, doi: 10.1109/ComPE49325.2020.9200031.
- [19] R. Gustriansyah, D. I. Sensuse, and A. Ramadhan, "Decision support system for inventory management in pharmacy using fuzzy analytic hierarchy process and sequential pattern analysis approach," *CONMEDIA 2015 - Int. Conf. New Media 2015*, 2016, doi: 10.1109/CONMEDIA.2015.7449153.