

Perbandingan Teknik Prediksi Pemakaian Obat Menggunakan Algoritma Simple Linear Regression dan Support Vector Regression

Sephia Pratista, Alwis Nazir*, Iwan Iskandar, Elvia Budianita, Iis Afrianty

Sains dan Teknologi, Teknik Informatika, Universitas Islam Negeri Sultan Syarif Kasim, Pekanbaru, Indonesia

Email: ¹11950121740@students.uin-suska.ac.id, ^{2,*}alwis.nazir@uin-suska.ac.id, ³iwan.iskandar@uin-suska.ac.id,

⁴elvia.budianita@uin-suska.ac.id, ⁵iis.afrianty@uin-suska.ac.id

Email Penulis Korespondensi: alwis.nazir@uin-suska.ac.id

Submitted: 09/09/2023; Accepted: 25/09/2023; Published: 27/09/2023

Abstrak—Puskesmas memiliki manajemen penting dalam memberikan pelayanan dan pengobatan kepada masyarakat, yakni manajemen obat. Untuk menghindari kesalahan pengelolaan ketersediaan stok obat, maka dilakukan prediksi. Prediksi dilakukan dengan membandingkan metode Data Mining menggunakan algoritma Simple Linear Regression dan metode Machine Learning menggunakan algoritma Support Vector Regression. Penelitian ini menggunakan data obat Paracetamol 500 mg dan Cetirizine dari bulan Januari 2020-Juni 2023. Dipilihnya algoritma ini karena variabel data bersifat kontinu dan memiliki rentang waktu, yakni 42 bulan (periode). Penelitian ini bertujuan untuk melihat seberapa besar kesalahan prediksi, dengan metode Mean Absolute Percentage Error (MAPE). Implementasi dilakukan menggunakan Bahasa pemrograman Python dengan pembagian data latihan dan uji sebesar 80%:20%. Berdasarkan percobaan yang telah dilakukan pada obat Paracetamol 500 mg menggunakan algoritma Simple Linear Regression, menghasilkan hasil MAPE sebesar 20.85% dengan kategori “Cukup” dan menggunakan algoritma Support Vector Regression menghasilkan MAPE sebesar 18.39% dengan kategori “Baik”. Percobaan pada obat Cetirizine menggunakan algoritma Simple Linear Regression menghasilkan hasil MAPE sebesar 18.39% dengan kategori “Baik” dan menggunakan algoritma Support Vector Regression menghasilkan MAPE sebesar 17.14% dengan kategori “Baik”. Berdasarkan hasil MAPE yang diperoleh, algoritma Support Vector Regression memiliki hasil prediksi yang lebih baik dibandingkan algoritma Simple Linear Regression.

Kata Kunci: Data Mining; Machine Learning; MAPE; Simple Linear Regression; Support Vector Regression

Abstract—Public Health Centers (Puskesmas) had a crucial role in furnishing society essential healthcare services and medication management. To preempt errors in stock management, a predictive approach is employed. This prediction methodology involves comparing Data Mining techniques utilizing the Simple Linear Regression algorithm and Machine Learning methodologies harnessing the Support Vector Regression algorithm. This research uses Paracetamol 500 mg and Cetirizine drug data from January 2020 to June 2023. The selection of these algorithms is motivated by the continuous nature of the data variables and their temporal span, spanning 42 months (period). The core aim of this study is to evaluate the magnitude of predictive errors using the Mean Absolute Percentage Error (MAPE) methodology. Implementing these methods was effectuated through the programming language Python with an 80%:20% partitioning of training and testing data. Drawing from experimental endeavors conducted concerning Paracetamol 500 mg, the utilization of the Simple Linear Regression algorithm, yields a MAPE score of 20.85%, categorized as 'Moderate,' whereas the application of the Support Vector Regression algorithm generates a MAPE of 18.39%, classified as 'Good.' Otherwise, experimentation on Cetirizine employing the Simple Linear Regression algorithm, employing an identical division of training and testing data, results in a MAPE of 18.39%, also classified as 'Good.' Meanwhile, resorting to the Support Vector Regression algorithm leads to a MAPE of 17.14%, falling under the 'Good' category. Based on the MAPE obtained, the Support Vector Regression algorithm has better prediction results than the Simple Linear Regression algorithm.

Keywords: Data Mining; Machine Learning; MAPE; Simple Linear Regression; Support Vector Regression

1. PENDAHULUAN

Puskesmas sebagai unit pelayanan kesehatan tingkat pertama dalam masyarakat memiliki manajemen dalam memberikan pelayanan dan pengobatan kepada masyarakat. Salah satu aspek penting dari manajemen ini adalah manajemen obat, karena puskesmas harus mengatur persediaan obat agar nantinya dapat memonitor sekaligus memprediksi obat-obatan untuk kedepannya.[1] Agar ketersediaan obat yang dibutuhkan di Puskesmas terjamin, diperlukan pengawasan terkait ketersediaan stok obat. Persediaan dan pemakaian obat tertuang dalam Laporan Penggunaan dan Permintaan Obat (LPLPO) yang memuat informasi mengenai persediaan dan penggunaan obat. Apabila permintaan obat terlalu banyak akan mengakibatkan surplus stok, sehingga akan mengakibatkan kedaluwarsa obat, sedangkan jenis obat lainnya memiliki permintaan yang tinggi tetapi stoknya sedikit, sehingga terjadi kekurangan obat untuk bulan berikutnya. [2], [3] Pada Puskesmas, pengelolaan data obat masih dilakukan secara manual, sehingga sangat memungkinkan terjadinya kesalahan dalam pengelolaan ketersediaan obat. Maka, untuk menghindari kesalahan tersebut dilakukanlah prediksi.

Prediksi dilakukan dengan membandingkan metode *Data Mining* menggunakan algoritma *Simple Linear Regression* (SLR) dan metode *Machine Learning* menggunakan algoritma *Support Vector Regression* (SVR). Data mining adalah suatu disiplin ilmu dalam ilmu komputer yang digunakan untuk tujuan menggali informasi tersembunyi dari sekumpulan data set untuk kemudian digunakan dalam pengambilan keputusan maupun kebijakan yang berguna dalam bidang-bidang tertentu.[4] Data mining merupakan proses pencarian dan pengenalan pola atau informasi dari sebuah data yang besar yang diperoleh dari berbagai macam basis data. Data mining menggunakan teknik statistik, matematika, kecerdasan buatan, pembelajaran mesin, dan visualisasi untuk mengekstrak, mengidentifikasi, serta

memecahkan masalah untuk mengetahui informasi yang baru berdasarkan data yang ada.[5]–[8] Machine Learning adalah bagian dari kecerdasan buatan (Artificial Intelligence) yang memfokuskan pada penggunaan data dan algoritma untuk memecahkan masalah dan membuat keputusan untuk menirukan cara manusia belajar dan meningkatkan kemampuan belajar komputer secara bertahap sehingga mampu menghasilkan sistem cerdas komputer tanpa penentuan aturan oleh manusia secara langsung yang dapat meningkatkan akurasi. Sistem dibuat agar mengenali pola kumpulan data yang digunakan sebagai data latih sehingga menghasilkan sebuah model yang bisa digunakan untuk melakukan prediksi suatu nilai (regresi). Machine learning memfokuskan diri pada pengembangan sistem yang memiliki kemampuan untuk belajar dan membuat keputusan secara mandiri, tanpa perlu diprogram ulang oleh manusia. Tidak hanya dapat menentukan aturan untuk tindakan yang optimal dalam pengambilan keputusan, mesin juga mampu beradaptasi terhadap perubahan yang terjadi.[9], [10]

Dipilihnya algoritma ini karena variabel data bersifat kontiniu dan memiliki rentang waktu, yakni selama 42 bulan (periode). Algoritma akan mencoba menemukan kurva untuk menemukan kecocokan antara vektor dan posisi kurva. Selanjutnya, dilakukan perhitungan nilai error terhadap kedua algoritma tersebut dengan metode *MAPE* (*Mean Absolute Percentage Error*). Hasil MAPE dari masing-masing algoritma kemudian digunakan sebagai acuan untuk membandingkan ketepatan hasil prediksi dari kedua algoritma tersebut. Penelitian ini bertujuan untuk melihat seberapa besar kesalahan prediksi bila dilakukan pencarian dengan menggunakan algoritma Simple Linear Regression dan Support Vector Regression. Diharapkan penelitian ini dapat membantu Puskesmas dalam mengelola stok obat dengan lebih baik dan memastikan ketersediaan obat memadai.

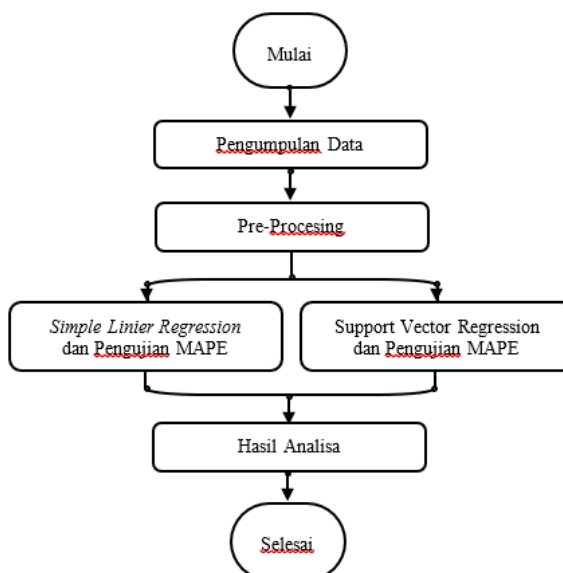
Terdapat beberapa penelitian yang menggunakan algoritma Support Vector Regression dan Linear Regression. Penelitian [11] melakukan prediksi beban menggunakan Support Vector Regression, penelitian menggunakan Linear Kernel yang dibandingkan dengan metode ANN dengan hasil MAPE untuk SVR dengan Linear Kernel sebesar 8,28% lebih besar dari metode ANN dengan MAPE 9,98%. Penelitian [12] melakukan prediksi arus lalu lintas jangka pendek menggunakan SVR dan dikombinasikan dengan KNN, sehingga menghasilkan hasil MAPE sebesar 14,726%. Penelitian [13] melakukan prediksi terhadap produksi padi menggunakan algoritma Simple Linear Regression, dengan nilai error dihitung menggunakan metode MAE dengan nilai 1,48950. Penelitian [14] melakukan prediksi terhadap persediaan obat jenis Tabetl menggunakan algoritma Simple Linear Regression dengan nilai akurasi sebesar 98,505%. Penelitian [15] melakukan prediksi penjualan menggunakan Linear Regression dan Metode Trend Moment, menghasilkan analisis hasil prediksi yang direkomendasikan adalah menggunakan metode Trend Moment dengan MAPE sebesar 0.439845%, sedangkan Linear Regression menghasilkan MAPE sebesar 1.511509%.

Keterbaruan dari penelitian ini adalah, penelitian dilakukan dengan membandingkan hasil nilai rata-rata absolut kesalahan (*error*) dari algoritma Simple Linear Regression dan Support Vector Regression, menggunakan bahasa pemrograman Python. Penelitian ini dilakukan karena dari berbagai sumber yang telah dibaca sebagai sumber acuan, belum ditemukan penelitian serupa mengenai perbandingan algoritma Simple Linear Regression dan Support Vector Regression pada prediksi pemakaian obat dengan menggunakan bahasa pemrograman Python.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Tahapan yang dilakukan dalam melakukan penelitian, ditampilkan dalam kerangka pada Gambar 1 berikut:



Gambar 1. Tahapan Penelitian



Pada tahap pengumpulan data, dipilih 2 macam obat yakni Paracetamol 500 mg dan Cetirizine dari bulan Januari 2020 hingga Juni 2023 (42 Bulan) dari salah satu Puskesmas di Kabupaten Kampar. Data obat yang telah dipilih akan ditampilkan dalam Tabel 1 dan Tabel 2 berikut:

Tabel 1. Data Obat Paracetamol 500 mg

BULAN	STOK AWAL	PENERIMAAN	PERSEDIAAN	PEMAKAIAN	SISA STOK
Jan-2020	8,858	5,000	13,858	1,091	12,767
Feb-2020	12,767		12,767	1,293	11,474
Mar-2020	11,474		11,474	1,495	9,979
...
Jun-2023	4,227		4,227	1,912	2,315

Tabel 2. Data Obat Cetirizine

BULAN	STOK AWAL	PENERIMAAN	PERSEDIAAN	PEMAKAIAN	SISA STOK
Jan-2020	1,914		1,914	1,220	694
Feb-2020	694	5000	5,694	957	4,737
Mar-2020	4,737		4,737	805	3,932
...
Jun-2023	1,907		1,907	1000	907

Selanjutnya, tahapan *pre-processing* untuk mendapatkan data yang diinginkan. Pada tahapan ini, dilakukan *data selection*, yaitu pemilihan atribut yang akan digunakan, yakni Bulan dan Pemakaian. Selanjutnya dilakukan *data cleaning*, yaitu pembersihan pada data yang terdapat *missing value*.

2.2 Simple Linear Regression (SLR)

Simple Linear Regression merupakan salah satu metode statistik. SLR digunakan untuk mengestimasi atau memprediksi hubungan antara dua variabel dalam penelitian kualitatif, dengan pendekatan untuk menetapkan hubungan antara satu variabel *dependen* dan satu variabel *independent*. Dengan asumsi bahwa hubungan antara variabel-variabel tersebut dapat didekati dengan persamaan garis lurus.[16] Persamaan umum pada metode Simple Linear Regression adalah sebagai berikut:

$$y = a + bX \tag{1}$$

Dengan :

a = Konstanta

b = Koefisien regresi

y = Variabel dependen (beabs)

X = Variabel independent (terikat)

Menghitung konstanta (a)

$$a = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2} \tag{2}$$

Menghitung koefisien (b)

$$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2} \tag{3}$$

2.3 Support Vector Regression

Support Vector Regression (SVR) merupakan algoritma yang menangani kasus regresi. SVR dapat mengatasi masalah *overfitting*. Tujuan dari SVR ini adalah untuk memetakan vektor input menggunakan fungsi kernel ke dimensi yang lebih tinggi. Dalam pemecahan kasus regresi, algoritma sekuensial untuk SVR mampu memberikan solusi dengan waktu komputasi yang lebih cepat dan lebih optimal apabila metode SVR konvensional sebagai pembandingnya.[17] Metode SVR digunakan untuk menyelesaikan masalah non-linear. SVR telah berhasil berhasil diterapkan dalam beberapa permasalahan dalam prediksi *time series*.

Tahapan implementasi SVR adalah sebagai berikut:

- Melakukan proses normalisasi data
- Menentukan kernel, parameter *kompleksitas* (C), dan *epsilon*
- Menerapkan algoritma SVR
- Melakukan pengujian nilai error

2.4 Mean Absolute Percentage Error (MAPE)

MAPE adalah indikator akurasi statistik. Indeks ini menunjukkan rata-rata persentase kesalahan absolut, semakin rendah MAPE, semakin tinggi akurasinya. MAPE merupakan rata-rata diferensiasi absolut antara nilai yang diprediksi



dan aktual. MAPE sangat intuitif interpretasi dalam hal kesalahan relative dan sering digunakan saat kuantitas untuk diprediksi diketahui tetap jauh di atas nol [18]–[20] MAPE menyatakan kesalahan dalam bentuk peresentasi dari hasil prediksi terhadap data actual. Persamaan untuk menghitung MAPE adalah sebagai berikut:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - y'_i}{y_i} \right| \times 100\% \quad (4)$$

Dengan:

n = jumlah data

y_i = nilai data latih

y'_i = nilai hasil prediksi

i = indeks

Kategori penentuan hasil MAPE dapat dilihat pada Tabel 3 berikut:

Tabel 3. Kategori Hasil MAPE

Hasil	Kategori
0% - 10%	Amat Baik
10% - 20%	Baik
20% - 50%	Cukup
50% - 100%	Buruk
> 100%	Amat Buruk

3. HASIL DAN PEMBAHASAN

3.1 Data yang Digunakan

Bedasarkan tahapan *pre-processing* yang telah dijelaskan sebelumnya, maka diperoleh dataset obat yang ditampilkan pada Tabel 4 dan Tabel 5 berikut:

Tabel 4. Dataset Obat Paracetamol 500 mg

Bulan	Pemakaian
Jan-2020	1,091
Feb-2020	1,293
Mar-2020	1,495
...	
Jun-2023	1,912

Tabel 5. Dataset Obat Cetirizine

Bulan	Pemakaian
Jan-2020	1,220
Feb-2020	957
Mar-2020	805
...	
Jun-2023	1,000

Atribut Bulan didefinisikan sebagai X (fitur) dan Pemakaian sebagai Y (target). Karena pada algoritma akan dilakukan perbandingan, maka akan dilakukan pembagian data latih dan uji sebesar 80%:20%. Dari jumlah data 42 bulan, sebanyak 33 bulan sebagai data latih, yakni dari bulan Januari 2020 hingga September 2022 dan 9 bulan data uji yakni dari bulan Oktober 2022 hingga Juni 2023.

Pada data obat Paracetamol 500 mg dan Cetirizine, akan diterapkan algoritma prediksi Simple Linear Regression dan Support Vector Regression. Kemudian, setelah dilakukan penerapan model, maka hasil prediksi bulanan dari tiap obat akan dihitung nilai rata-rata kesalahan absolutnya dalam bentuk persentase (MAPE). Hasil MAPE akan digunakan untuk menentukan kinerja kedua algoritma yang digunakan pada kedua data obat ini.

3.2 Implementasi dan Pengujian Algoritma Simple Linear Regression

Penerapan algoritma Linear Regression pada Python

```
regressor = LinearRegression()
regressor.fit(X_train, y_train)
```

Code regressor akan menerapkan model Linear Regression dan *regressor.fit(X_train, y_train)* akan mengajarkan model untuk melakukan pelatihan data.

Setelah algoritma diterapkan, dilakukan pencarian intercept, coefficient, score untuk melakukan prediksi

```
model = LinearRegression().fit(X, y)
a = model.intercept_
b = model.coef_
r = model.score(X, y)
```

Code model = LinearRegression().fit(X, y) akan mempelajari hubungan antara fitur X dan target Y. *Code* ini juga digunakan untuk menentukan nilai *intercept*, *coefisien*, dan *score*.

Kemudian lakukan prediksi

```
y_pred = regressor.predict(X_test)
```

Code y_pred = regressor.predict(X_test) akan melakukan prediksi berdasarkan model pelatihan data dengan memberikan data pengujian sebagai input ke model.

Setelah hasil prediksi diperoleh, kemudian lakukan perhitungan nilai error dengan metode MAPE

```
mape = mean_absolute_percentage_error(y_test, y_pred)
print("MAPE: {:.2f}%".format(mape * 100))
```

Code mape = mean_absolute_percentage_error(y_test, y_pred) akan menghitung persentasi rata rata kesalahan absolut berdasarkan target pengujian dan target prediksi.

3.3 Implementasi dan Pengujian Algoritma Support Vector Regression

Sebelum melakukan pengujian terhadap algoritma SVR, terlebih dahulu lakukan normalisasi dataset sebagai pendekatan sistematis yang meminimalkan pengulangan atau penghilangan pada data yang tidak perlu, sehingga database dapat bekerja secara optimal. Tujuan dilakukannya normalisasi adalah mengubah nilai data menjadi lebih mudah dipahami agar rentang pada nilai tidak terlalu besar dan unuk menghindari terjadinya anomali, *overfitting*, serta ketidakkonsistenan data. Normalisasi yang umum digunakan adalah *min-max-scaler*.

Penerapan normalisasi *min-max-scaler*

```
scaler_X = MinMaxScaler()
scaler_y = MinMaxScaler()
X_train_scaled = scaler_X.fit_transform(X_train)
y_train_scaled = scaler_y.fit_transform(y_train.reshape(-1, 1)).flatten()
```

Code ini akan menerapkan pada atribut X (fitur) dan Y (target). Kemudian dilakukan penghitungan parameter penskalaan berdasarkan data latih dan mengubah data latih sesuai dengan skala yang telah diterapkan

Menentukan dan menerapkan metode SVR

```
svr = SVR(kernel='rbf', C=1.0, epsilon=0.1)
svr.fit(X_train_scaled, y_train_scaled)
```

Code svr = SVR(kernel='rbf', C=1.0, epsilon=0.1) akan membuat model algoritma SVR dengan menggunakan *kernel RBF*, *parameter complexity 1.0*, dan *epsiolon 0.1*. Kemudian model diterapkan pada data latih atribut X dan Y yang telah diskalakan.

Setelah metode SVR diterapkan, lakukan prediksi

```
y_pred = scaler_y.inverse_transform(y_pred_scaled.reshape(-1, 1)).flatten()
```

Code scaler_y.inverse_transform(y_pred_scaled.reshape(-1, 1)) akan melakukan prediksi menggunakan *scaler_y* yang telah dibuat sebelumnya. Kemudian *inverse_transform*, dilakukan pembalikan nilai yang skalanya telah diubah menjadi nilai aslinya. Lalu *reshape(-1, 1)* digunakan untuk memastikan kembali matriks berbentuk satu kolom. *Code flatten()* digunakan untuk mengubah nilai menjadi vektor.

Setelah hasil prediksi diperoleh, kemudian lakukan perhitungan nilai error dengan metode MAPE

```
mape = mean_absolute_percentage_error(y_test, y_pred)
print("MAPE: {:.2f}%".format(mape * 100))
```

Code $mape = mean_absolute_percentage_error(y_test, y_pred)$ akan menghitung persentasi rata rata kesalahan absolut berdasarkan target pengujian dan target prediksi.

3.4 Pembahasan

3.4.1 Hasil Prediksi

Hasil prediksi dengan menggunakan algoritma Simple Linear Regression dan Support Vector Regression untuk obat Paracetamol 500 mg akan ditampilkan dalam Tabel 6 dan Cetirizine akan ditampilkan dalam Tabel 7.

a. Untuk Obat Paracetamol 500 mg

Tabel 6. Tabel Hasil Prediksi Obat Paracetamol 500 mg

Bulan	Periode	Aktual	Hasil Prediksi Obat Paracetamol 500 mg	
			Simple Linear Regression	Support Vector Regression
Okt-2022	34	1800	2046.280303	1715.062166
Nov-2022	35	2216	2053.038324	1713.750594
Des-2022	36	1887	2059.796346	1714.193051
Jan-2023	37	1,443	2066.554367	1714.540370
Feb-2023	38	1,706	2073.312389	1713.782003
Mar-2203	39	1,638	2080.070410	1711.662731
Apr-2023	40	1,474	2086.828431	1708.474409
Mei-2023	41	1,831	2093.586453	1704.801490
Jun-2023	41	1,912	2100.344474	1701.286755

b. Untuk Obat Cetirizine

Tabel 7. Hasil Prediksi Obat Cetirizine

Bulan	Periode	Aktual	Hasil Prediksi Obat Cetirizine	
			Simple Linear Regression	Support Vector Regression
Okt-2022	34	1800	1081.399621	1122.041048
Nov-2022	35	1300	1082.298351	1119.450928
Des-2022	36	1,000	1083.197081	1118.449233
Jan-2023	37	1,300	1084.095811	1118.549112
Feb-2023	38	1,150	1084.994541	1119.157179
Mar-2023	39	1,375	1085.893271	1119.710904
Apr-2023	40	1,600	1086.792001	1119.774991
Mei-2023	41	1,300	1087.690731	1119.090116
Jun-2023	42	1,000	1088.589461	1117.577835

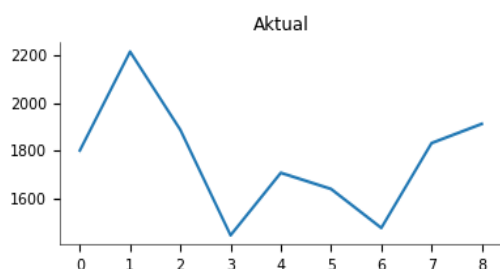
3.4.2 Grafik Hasil Prediksi

Berdasarkan prediksi yang telah dilakukan, maka dapat dilihat grafiknya sebagai berikut:

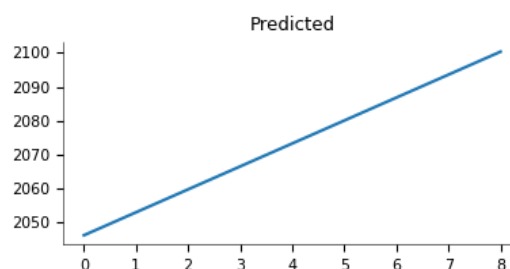
a. Untuk obat Paracetamol 500 mg

1. Menggunakan Simple Linear Regression (SLR)

Grafik data aktual dan grafik prediksi obat Paracetamol 500 mg menggunakan algoritma SLR akan ditampilkan pada Grafik 1 dan Grafik 2 berikut:



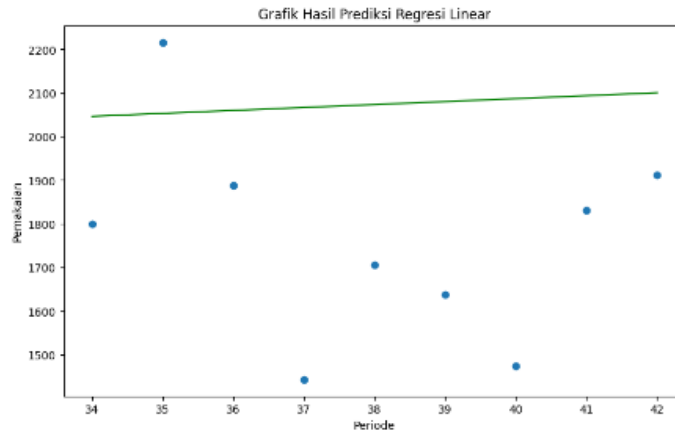
Grafik 1. Data Aktual Obat Paracetamol 500 mg - SLR



Grafik 2. Hasil Prediksi Obat Paracetamol 500 mg - SLR



Grafik 1 menunjukkan pola data aktual yang akan diprediksi. Dapat dilihat bahwa data aktual memiliki pola yang bersifat naik-turun. Grafik 2 menunjukkan garis regresi hasil prediksi yang cenderung naik. Selanjutnya, rata-rata hasil prediksi hasil prediksi obat Paracetamol 500 mg terhadap bulan prediksi menggunakan algoritma SLR akan ditampilkan pada Grafik 3 berikut:

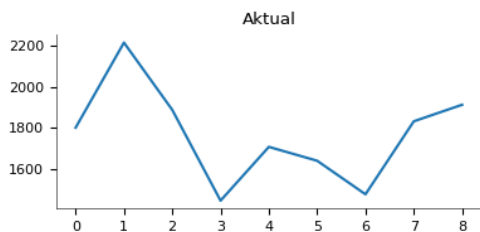


Grafik 3. Hasil Prediksi Terhadap Bulan Obat Paracetamol 500 mg - SLR

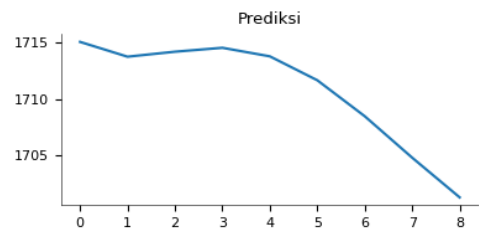
Grafik 3 menunjukkan hasil prediksi untuk 9 bulan selanjutnya menunjukkan kenaikan, namun data aktual dan hasil prediksi berada pada rentang yang jauh.

2. Menggunakan Support Vector Regression (SVR)

Grafik data aktual dan grafik prediksi obat Paracetamol 500 mg menggunakan algoritma SVR akan ditampilkan pada Grafik 4 dan Grafik 5 berikut:

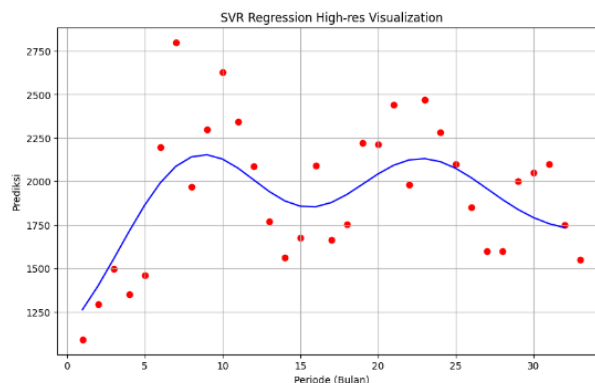


Grafik 4. Data Aktual Obat Paracetamol 500 mg - SVR



Grafik 5. Hasil Prediksi Paracetamol 500 mg - SVR

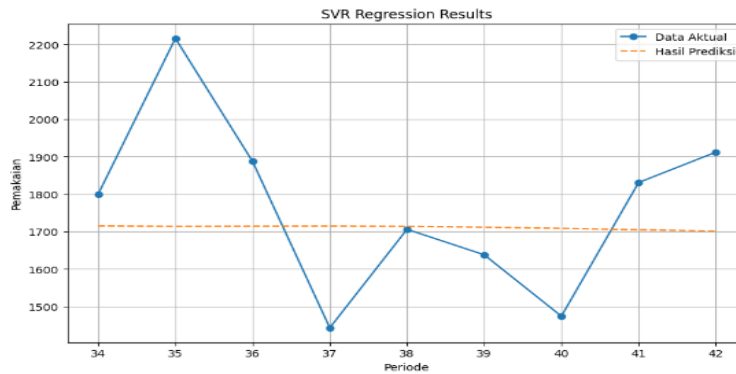
Grafik 4 menunjukkan pola data aktual yang akan diprediksi. Dapat dilihat bahwa data aktual memiliki pola yang bersifat naik-turun. Grafik 5 menunjukkan garis vektor regresi hasil prediksi yang cenderung turun. Hasil dari model SVR pada obat Paracetamol 500 mg akan ditampilkan dalam grafik visualisasi tingkat tinggi pada Grafik 6 berikut:



Grafik 6. Visualisasi Paracetamol 500 mg

Grafik 6 visualisasi SVR dengan titik merah menunjukkan *data-points* dan garis biru menunjukkan garis *hyperlane non-linear*. Dapat dilihat bahwa titik merah paling kanan dan paling kiri tidak terhubung dengan hyperlane. Hal ini disebabkan karena hyperlane dari SVR menganggap titik tersebut sebagai outlier karena jaraknya terhadap data-points yang lain terlalu jauh. Selain itu, titik yang berada diluar margin hyperlane juga dianggap sebagai outlier karena telah melewati batas margin yang dilalui hyperlane.

Selanjutnya, rata-rata hasil prediksi hasil prediksi obat Paracetamol 500 mg terhadap bulan prediksi menggunakan algoritma SVR akan ditampilkan pada Grafik 7 berikut:



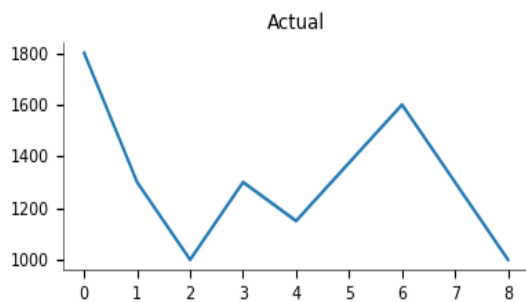
Gtapak 7. Hasil Prediksi Terhadap Bulan Obat Paracetamol 500 mg - SVR

Grafik 7 menunjukkan hasil prediksi untuk 9 bulan selanjutnya menunjukkan penurunan yang tidak terlalu signifikan. Namun, garis regresi terhadap beberapa data aktual memiliki jarak yang jauh.

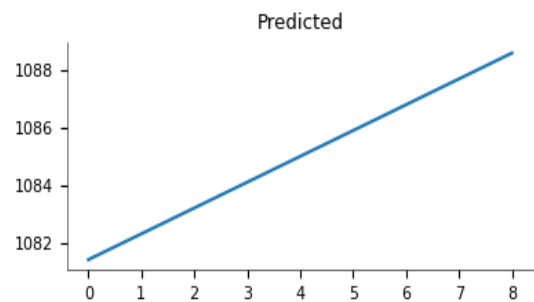
b. Untuk obat Cetirizine

1. Menggunakan Simple Linear Regression (SLR)

Grafik data aktual dan grafik prediksi obat Cetirizine menggunakan algoritma SLR akan ditampilkan pada Grafik 8 dan Grafik 9 berikut:

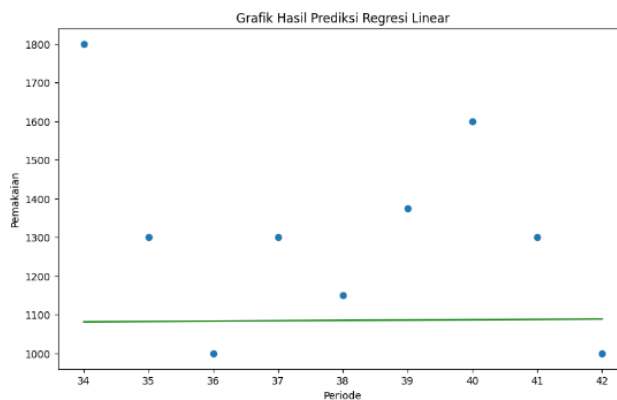


Grafik 8. Data Aktual Obat Cetirizine - SLR



Grafik 9 Hasil Prediksi Obat Cetirizine - SLR

Grafik 8 menunjukkan pola data aktual yang akan diprediksi. Dapat dilihat bahwa data aktual memiliki pola yang bersifat naik-turun. Grafik 9 menunjukkan garis regresi hasil prediksi yang cenderung naik. Selanjutnya, rata-rata hasil prediksi hasil prediksi obat Cetirizine terhadap bulan prediksi menggunakan algoritma SLR akan ditampilkan pada Grafik 10 berikut:

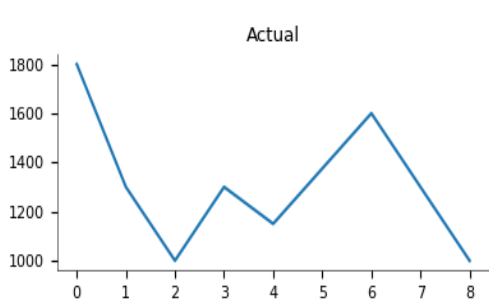


Grafik 10. Hasil Prediksi Terhadap Bulan Obat Cetirizine - SLR

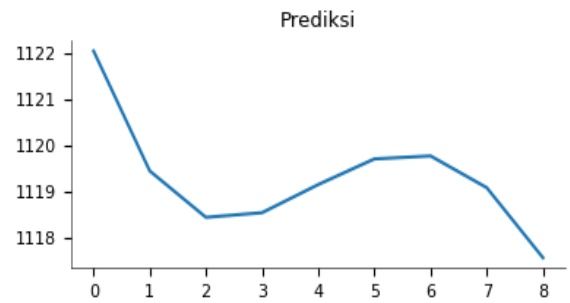
Grafik 10 menunjukkan hasil prediksi untuk 9 bulan selanjutnya menunjukkan kenaikan yang tidak begitu signifikan, namun data aktual dan hasil prediksi berada pada rentang yang jauh.

2. Menggunakan Support Vector Regression (SVR)

Grafik data aktual dan grafik prediksi obat Cetirizine menggunakan algoritma SVR akan ditampilkan pada Grafik 11 dan Grafik 12 berikut:

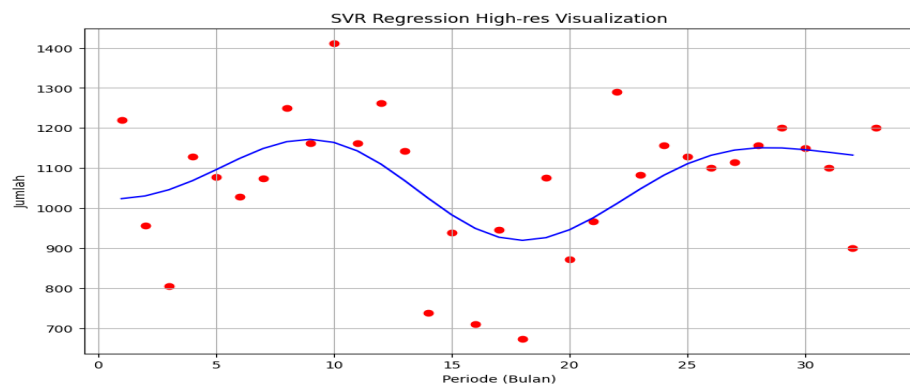


Grafik 11. Data Aktual Obat Cetirizine - SVR



Grafik 12. Hasil Prediksi Obat Cetirizine - SVR

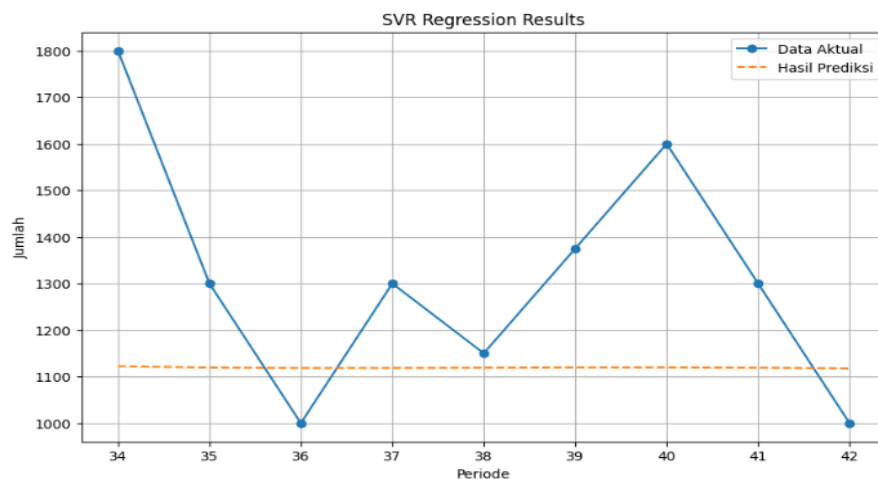
Grafik 11 menunjukkan pola data aktual yang akan diprediksi. Dapat dilihat bahwa data aktual memiliki pola yang bersifat naik-turun. Grafik 12 menunjukkan garis vektor regresi hasil prediksi yang cenderung turun. Hasil dari model SVR pada obat Cetirizine akan ditampilkan dalam grafik visualisasi tingkat tinggi pada Grafik 13 berikut:



Grafik 13. Visualisasi Cetirizine

Grafik 13 visualisasi SVR dengan titik merah menunjukkan *data-points* dan garis biru menunjukkan garis *hyperlane non-linear*. Dapat dilihat bahwa titik merah paling kanan tidak terhubung dengan hyperlane. Hal ini disebabkan karena hyperlane dari SVR menganggap titik tersebut sebagai *outlier* karena jaraknya terhadap data-points yang lain terlalu jauh. Selain itu, titik yang berada diluar margin hyperlane juga dianggap sebagai outlier karena telah melewati batas margin yang dilalui hyperlane.

Selanjutnya, rata-rata hasil prediksi hasil prediksi obat Cetirizine terhadap bulan prediksi menggunakan algoritma SVR akan ditampilkan pada Grafik 14 berikut:



Grafik 14. Hasil Prediksi Cetirizine Terhadap Bulan - SVR

Grafik 14 menunjukkan hasil prediksi untuk 9 bulan selanjutnya menunjukkan penurunan yang tidak terlalu signifikan. Namun, garis regresi terhadap beberapa data aktual memiliki jarak yang jauh.

3.4.3 Hasil Mean Absolute Percentage Error (MAPE)

Berdasarkan prediksi yang dilakukan, maka diperoleh hasil kesalahan error sebagai berikut:

Tabel 8. Hasil Mean Absolute Percentage Error

Algoritma	Nama Obat	Hasil	Kategori
Simple Linear Regression	Paracetamol 500 mg	20.85%	Cukup
	Cetirizine	18.39%	Baik
Support Vector Regression	Paracetamol 500 mg	10.46%	Baik
	Cetirizine	17.14%	Baik

Pada obat Paracetamol 500 mg, apabila diprediksi menggunakan algoritma Simple Linear Regression (SLR) diperoleh hasil MAPE sebesar 20.85% dengan kategori “Cukup” sedangkan apabila diprediksi menggunakan algoritma Support Vector Regression (SVR) diperoleh hasil MAPE sebesar 10.46% dengan kategori “Baik”. Perbedaan ini dapat terjadi karena pola data yang naik turun dengan sangat signifikan. Pada pertengahan tahun 2020, data pemakaian obat yang awalnya berada pada kisaran 1400-an, tiba-tiba melonjak hingga ke angka 2000-an. Kemudian, pada pertengahan tahun 2022 mulai mengalami penurunan pemakaian kembali hingga ke kisaran angka 1400-an. Pola data naik-turun yang rumit akan mempersulit pembelajaran data bagi algoritma SLR yang mencoba untuk menemukan garis linear terbaik yang paling mendekati pola data. Namun, prediksi dapat dilakukan dengan baik oleh algoritma SVR karena algoritma ini dapat menemukan fungsi regresi yang paling mendekati data dengan juga mempertimbangkan margin disekitar garis regresi, sehingga dapat menangani data yang memiliki pola non-linear.

Kemudian, pada obat Cetirizine apabila diprediksi menggunakan algoritma Simple Linear Regression (SLR) diperoleh hasil MAPE sebesar 18.39% dengan kategori “Baik” sedangkan apabila diprediksi menggunakan algoritma Support Vector Regression (SVR) diperoleh hasil MAPE sebesar 17.14% dengan kategori “Baik”. Walaupun kedua algoritma menunjukkan kinerja yang baik, namun algoritma SVR sedikit lebih unggul. Hal ini terjadi karena pola data yang naik turun, namun tidak begitu signifikan, sehingga kedua algoritma dapat menemukan fungsi regresi yang mendekati pola data.

4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan pada obat Paracetamol 500 mg, apabila diprediksi menggunakan algoritma Simple Linear Regression diperoleh hasil MAPE sebesar 20.85% dengan kategori “Cukup” sedangkan apabila diprediksi menggunakan algoritma Support Vector Regression, menggunakan kernel RBF, dengan parameter *complexity* 1.0, dan parameter epsilon 0.1, diperoleh hasil MAPE sebesar 10.46% dengan kategori “Baik”. Kemudian Pada obat Cetirizine, dengan pembagian data uji dan latih yang sama, apabila diprediksi menggunakan algoritma Simple Linear Regression diperoleh hasil MAPE sebesar 18.39% dengan kategori “Baik” sedangkan apabila diprediksi menggunakan algoritma Support Vector Regression dengan parameter yang sama diperoleh hasil MAPE sebesar 17.14% dengan kategori “Baik”. Berdasarkan indikator MAPE, dapat diketahui semakin rendah hasil MAPE maka semakin tinggi akurasi. Sebaliknya, semakin tinggi hasil MAPE maka semakin rendah akurasi. Dengan demikian, dalam penelitian ini disimpulkan algoritma Support Vector Regression adalah algoritma terbaik dalam melakukan prediksi dibandingkan dengan algoritma Simple Linear Regression.

REFERENCES

- [1] Dahlia dan Andri, “Implementasi Data Mining untuk Prediksi Persediaan Obat pada Puskesmas Kertapati menggunakan Regresi Linier Berganda,” JURNAL SISTEM DAN INFORMATIKA, hlm. 95–103, Nov 2020, doi: 10.30864/jsi.v15i2.331.
- [2] D. Suwardiyanto, M. Nur Shodiq, D. Hidayat Kusuma, dan T. Oktalita Sari, “Sistem Prediksi Kebutuhan Obat di Puskesmas Menggunakan Metode Least Square,” Jurnal Informatika: Jurnal Pengembangan IT, vol. 4, no. 1, hlm. 75–80, Jan 2019, doi: 10.30591/jpit.v4i1.1085.
- [3] D. Abdianto, Elisawati, F. Tawakal, dan Masrizal, “Prediksi Stok Obat Menggunakan Metode Learning Vector Quantization Studi Kasus Puskesmas Dumai Barat,” Prosiding Seminar Sains Nasional dan Teknologi, vol. 1, no. 1, hlm. 68–74, 2021.
- [4] H. Noor, A. Dharmawati, dan T. Wahyu Qur, “Penerapan Algoritma K-Means Clustering Analysis Pada Kasus Penderita Hiv/Aids (Studi Kasus Kabupaten Banjar),” Technologia, vol. 12, no. 2, hlm. 72–76, Apr 2021.
- [5] Y. Andini, J. Tata Hardinata, dan Y. Pranayama Purba, “Penerapan Data Mining Terhadap Tata Letak Buku Di Perpustakaan Sintong Bingei Pematangsiantar Menggunakan Metode Apriori,” Jurnal Times, vol. XI, no. 1, hlm. 9–15, 2022, [Daring]. Tersedia pada: <http://ejournal.stmik-time.ac.id>
- [6] S. P. Dewi, N. Nurwati, dan E. Rahayu, “Penerapan Data Mining Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode K-Nearest Neighbor,” Building of Informatics, Technology and Science (BITS), vol. 3, no. 4, hlm. 639–648, Mar 2022, doi: 10.47065/bits.v3i4.1408.
- [7] N. Karolina, “Data Mining Pengelompokan Pasien Rawat Inap Peserta BPJS Menggunakan Metode Clustering (Studi Kasus : RSU.Bangkalan),” JOURNAL OF INFORMATION AND TECHNOLOGY UNIMOR, vol. 1, no. 2, hlm. 47–53, Sep 2021, [Daring]. Tersedia pada: www.kaputama.ac.id
- [8] Haryan, D. Nofriansyah, dan I. Mariami, “Implementasi Data Mining Untuk Pengelempokan Buku Di Perpustakaan Yayasan Nurul Islam Indonesia Baru Dengan Metode K-Means Clustering,” Jurnal CyberTech, vol. 1, no. 1, hlm. 1–12, Sep 2021, [Daring]. Tersedia pada: <https://ojs.trigunadharma.ac.id/index.php/jct/index>
- [9] M. Syauqi Haris, A. Naseh Khudori, dan W. Teja Kusuma, “Perbandingan Metode Supervised Machine Learning Untuk Prediksi Prevalensi Stunting Di Provinsi Jawa Timur,” Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK), vol. 9, no. 7, hlm. 1571–1576, Des 2022, doi: 10.25126/jtiik.202296744.



- [10] R. Pratama, M. I. Herdiansyah, D. Syamsuar, dan A. Syazili, “Prediksi Customer Retention Perusahaan Asuransi Menggunakan Machine Learning,” *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, vol. 12, no. 1, hlm. 96–104, Mar 2023, doi: 10.32736/sisfokom.v12i1.1507.
- [11] L. W. Chong, D. Rengasamy, Y. W. Wong, dan R. K. Rajkumar, “Load Prediction Using Support Vector Regression,” dalam *TENCON 2017 - 2017 IEEE Region 10 Conference, IEEE*, Nov 2017, hlm. 1069–1074. doi: 10.1109/TENCON.2017.8228016.
- [12] G. Lin, A. Lin, dan D. Gu, “Using Support Vector Regression and K-Nearest Neighbors For Short-Term Traffic Flow Prediction Based On Maximal Information Coefficient,” *Inf Sci (N Y)*, vol. 608, hlm. 517–531, Agu 2022, doi: 10.1016/j.ins.2022.06.090.
- [13] A. A. Suryanto dan A. Muqtadir, “Penerapan Metode Mean Absolute Error (MEA) Dalam Algoritma Regresi Linear Untuk Prediksi Produksi Padi,” *SAINTEKBU: Jurnal Sains dan Teknologi*, vol. 11, no. 1, hlm. 78–83, 2019.
- [14] Harsiti, Z. Muttaqin, dan E. Srihartini, “Penerapan Metode Regresi Linier Sederhana Untuk Prediksi Persediaan Obat Jenis Tablet,” *Jurnal Sistem Informasi*, vol. 9, no. 1, hlm. 12–16, Mar 2022.
- [15] A. Suhaidi Batubara, H. Dafitri, dan I. Faisal, “Analysis Of Linear Regression And Trend Moment Methods In Predicting Sales Using Mape,” *Jurnal Sistem Informasi dan Ilmu Komputer Prima*, vol. 6, no. 1, hlm. 75–81, 2022, doi: <https://doi.org/10.34012/jurnalsisteminformasidanilmukomputer.v6i1.2919>.
- [16] T. Wahyudi dan D. Septya Arroufu, “Implementation Of Data Mining Prediction Delivery Time Using Linear Regression Algorithm,” *Journal of Applied Engineering and Technological Science*, vol. 4, no. 1, hlm. 84–92, 2022, doi: <https://doi.org/10.37385/jaets.v4i1.918>.
- [17] Nendi dan A. Wibowo, “Prediksi Jumlah Pengiriman Barang Menggunakan Kombinasi Metode Support Vector Regression, Algoritma Genetika dan Multivariate Adaptive Regression Splines,” *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, vol. 7, no. 6, hlm. 1169–1176, Des 2020, doi: 10.25126/jtiik.202072441.
- [18] D. Iskandar, W. Adi, dan S. Wibowo, “Data Mining Dalam Prediksi Jumlah Pasien Dengan Regresi Linear Dan Exponential Smoothing,” *Jurnal Sistem Informasi dan Sains Teknologi*, vol. 5, no. 1, hlm. 1–8, 2023.
- [19] R. Novita, I. Yani, dan G. Ali, “Sistem Prediksi untuk Penentuan Jumlah Pemesanan Obat Menggunakan Regresi Linier,” *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 2, no. 1, hlm. 62–70, Apr 2022.
- [20] A. Anggrawan, N. Azmi, U. Bumigora, dan I. Anthonyangrawan, “Prediksi Penjualan Produk Unilever Menggunakan Metode Regresi Linear Sales Prediction of Unilever Products using the Linear Regression Method,” *Jurnal Bumigora Information Technology (BITe)*, vol. 4, no. 2, hlm. 123–132, Nov 2022, doi: 10.30812/bite.v4i2.2416.