



PENERAPAN METODE REGRESI LINIER UNTUK PREDIKSI PENYAKIT DIARE MENGGUNAKAN PYTHON (STUDI KASUS : UPT PUSKESMAS TERAWAS)

Denci Puspa A¹, Andri Anto T.S², Lukman Sunardi³

^{1,2,3}Program Studi Informatika, Universitas Bina Insan, Lubuklinggau

e-mail: *¹enciaprilia04@gmail.com, ²andri@gmail.com, ³lukmansunardi@univbinainsan.ac.id

Abstrak

Penyakit diare merupakan penyakit yang paling banyak dijumpai kasusnya. Hal ini ditunjukkan dengan angka penderita diare di Puskesmas di wilayah Kabupaten/Kota yang tinggi setiap tahunnya menempati urutan tertinggi dan merupakan penyakit yang paling banyak dijumpai kasusnya. Hal ini ditunjukkan dengan angka penderita diare di Puskesmas di wilayah Kota/Kabupaten, Namun sulit untuk mengetahui jumlah pasien penderita diare yang sesungguhnya karena mengingat banyaknya penderita yang tidak terdata karena tidak mengunjungi tempat – tempat pelayanan kesehatan. Melihat kondisi, untuk itu diperlukan kemampuan untuk menentukan faktor yang paling dominan dalam memprediksi jumlah pasien penyakit diare yang ada di UPT Puskesmas Terawas. Regresi Linier digunakan untuk memprediksi hubungan antara variabel terikat (Y) dan variabel bebas (X). Adanya sistem prediksi jumlah pasien penyakit diare, maka UPT Puskesmas Terawas dapat memprediksi jumlah pasien penyakit diare di masa mendatang, sehingga dapat mempersiapkan segala kebutuhan dan tindakan untuk menangani penyakit diare di UPT Puskesmas Terawas.

Kata kunci—3-5 kata kunci; Penyakit Diare, Prediksi, Regresi Linier Berganda

Abstract

Diarrhea is the disease with the most cases. This is shown by the high number of diarrhea sufferers in Community Health Centers in Regency/City areas which ranks highest every year and is the disease with the most cases. This is shown by the number of diarrhea sufferers in Community Health Centers in the City/Regency area. However, it is difficult to know the actual number of patients suffering from diarrhea because we remember that many sufferers are not recorded because they do not visit health service places. Looking at the conditions, it is necessary to be able to determine the most dominant factors in predicting the number of patients with diarrhea at the Terawas Health Center UPT. Linear Regression is used to predict the relationship between a dependent variable (Y) and a series of independent variables (X). With a prediction system for the number of patients with diarrhea, the UPT Puskesmas Terawas can predict the number of patients with diarrhea in the future, so that it can prepare all the needs and actions to deal with diarrhea at the UPT Puskesmas Terawas.

Keywords—3-5 keywords: *Diarrhea, Prediction, Multiple Linear Regression*

I. PENDAHULUAN

Permasalahan penyakit merupakan salah satu faktor kunci dalam upaya peningkatan kesehatan masyarakat,

mengingat jumlah orang yang terjangkit penyakit terus meningkat dan menyebabkan kematian. Menurut data WHO (Organisasi Kesehatan Dunia), diare merupakan penyebab utama kematian pada anak. Di



Indonesia, diare merupakan penyakit dasar yang disebabkan oleh kontaminasi makanan, penyakit bakteri, gula palsu, fruktosa, dan lain-lain. Diare dapat terjadi karena higienitas dan disinfeksi yang buruk, kondisi yang tidak sehat, iklim yang padat, dan fasilitas klinik yang kurang mendukung. Diare sendiri disebabkan oleh infeksi dan mikroba, termasuk virus, bakteri, dan parasit (Santos, 2019). Hal ini ditunjukkan dengan angka penderita diare di Puskesmas di wilayah Kabupaten/Kota yang tinggi setiap tahunnya menempati urutan tertinggi dan merupakan penyakit yang paling banyak dijumpai kasusnya (Penyakit et al., 2019). Melihat kondisi, untuk itu diperlukan kemampuan untuk menentukan faktor yang paling dominan dalam memprediksi jumlah pasien penyakit diare yang ada di UPT Puskesmas Terawas dan hasil dari penelitian ini diharapkan dapat berfungsi dengan optimal untuk kasus prediksi yang terjadi (Linier et al., 2023).

Penelitian ini menggunakan beberapa referensi penelitian yang dilakukan oleh peneliti sebelumnya. Penelitian terdahulu pertama yang berjudul Implementasi Data Mining Dengan Algoritma Multiple Linear Regression Untuk Memprediksi Penyakit Diabetes yang telah menghasilkan bahwa metode multiple linear regression untuk memprediksi penyakit diabetes, serta melakukan evaluasi menggunakan RMSE (root mean square error) dan menghasilkan nilai RMSE sebesar 0.403 (Normah et al., 2022). Penelitian terdahulu kedua yang telah menghasilkan bahwa atribut (usia, nyeri dada dan hasil test) berpengaruh signifikan terhadap gejala penyakit (Wijayadhi et al., 2023). Penelitian terdahulu ketiga Hasil dari penelitian mengungkapkan bahwa proses implementasi machine learning terhadap prediksi tingkat kasus penyakit dengan algoritma klasifikasi dengan teknik decision tree menghasilkan akurasi yang sangat baik

(Data et al., 2023). Penelitian terdahulu ke empat, Hasil analisis menunjukkan Kanker paru-paru adalah salah satu jenis kanker yang paling mematikan di seluruh dunia. Oleh karena itu, upaya untuk memprediksi kemungkinan terkena kanker paru-paru sangat penting dalam pencegahan dan penanganannya. Salah satu cara untuk memprediksi kemungkinan terkena kanker paru-paru adalah dengan menggunakan algoritma regresi linier. Hasil analisis menunjukkan bahwa algoritma regresi linier dapat digunakan untuk memprediksi kemungkinan terkena kanker paru-paru dengan akurasi sekitar 90% dan mampu memberikan hasil yang baik dengan nilai Root Mean Squared Error: 0.686 +/- 0.000 dan Squared Error: 0.471 +/- 0.546 (Wahid et al., 2023). Penelitian terdahulu ke lima, dengan hasil analisis Balita merupakan kelompok paling rentan terhadap masalah gizi apabila ditinjau dari masalah kesehatan dan gizi, kejadian balita gizi buruk di kabupaten Bojonegoro untuk persentase balita gizi buruk perkecamatan. variabel yang berpengaruh signifikan terhadap balita yang mengalami gizi buruk, yaitu persentase balita ditimbang empat kali atau lebih dalam empat bulan terakhir sebesar 2,117 dan persentase balita kurus mendapatkan makanan tambahan sebesar - 0,438 dengan akurasi model regresi yang diperoleh R-Square 74,3% (Yuanita & Yuanita, 2023).

Data Mining adalah analisis data observasi dalam jumlah besar untuk menemukan hubungan yang sebelumnya tidak diketahui dan dua metode baru untuk merangkum data sehingga dapat dimengerti dan berguna bagi orang yang memilih data. Data mining didefinisikan sebagai proses menemukan hubungan, pola dan tren yang baru dan bermakna dengan menyaring data dalam jumlah yang sangat besar, disimpan dalam memori, menggunakan teknik Teknik pengenalan pola sebagai teknik

statistik dan matematika (Simbolon, 2021). Menurut kamus besar bahasa Indonesia, prediksi adalah hasil dari kegiatan memprediksi atau meramal atau memperkirakan nilai pada masa yang akan datang dengan menggunakan data masa lalu [12] (Kafil & Industri, 2019). Prediksi bisa bersifat kualitatif (tidak berbentuk angka) maupun kuantitatif (berbentuk angka) (Syafuruddin et al., n.d.). Regresi linier adalah suatu metode statistik yang digunakan untuk membentuk model atau hubungan antara satu atau lebih variabel bebas X dan variabel respon Y. Analisis regresi dengan satu variabel bebas X disebut sebagai regresi linier sederhana, sedangkan jika terdapat lebih dari satu variabel bebas X disebut sebagai regresi linier berganda (Syilfi, dwi ispriyanti, 2017). Regresi Linier merupakan salah satu metode yang optimal untuk sistem prediksi dengan menghubungkan antara variabel terikat (Y) dan variabel bebas (X). Tujuan dari metode ini adalah untuk memprediksi nilai Y untuk nilai X yang diberikan (Shidiq et al., 2020). Hal ini membuat anggapan bahwa hubungan antara dua variabel tersebut adalah linier. Regresi linier langsung memutuskan hubungan antara variabel dependen dan variabel independent secara bebas dengan menentukan garis terbaik (Garis Regresi) (Hermawan & Nugroho, 2023). Supaya memiliki sistem dan metode yang dapat digunakan untuk memprediksi jumlah pasien penyakit kedepannya, salah satunya untuk persiapan stok obat, infus, dan sebagainya untuk yang akan datang (Wahid et al., 2023). Diare adalah penyakit yang dapat mengakibatkan penurunan nafsu makan, sakit perut, rasa lelah, hingga penurunan berat badan (Ibrahim & Sartika, 2021).

Diare juga dapat mengakibatkan kehilangan cairan elektrolit secara mendadak sehingga mengakibatkan penderita mengalami komplikasi seperti

dehidrasi, kerusakan organ, bahkan yang lebih parah nya lagi yaitu koma.

Bahasa Pemrograman Python diciptakan oleh Guido Van Rossum di Belanda pada tahun 1990 dan namanya diambil dari acara TV favorit Guido Monty, Python's Flyin Circus (Gede et al., 2022). Adapun library python, yaitu:

1. Pandas: manipulasi dan analisis data
2. NumPy: komputasi numerik dengan Python
3. Matplotlib: plot yang memungkinkan membuat statis, animasi, dan interaktif visualisasi dengan Python.
4. Seaborn: visualisasi data statistik berdasarkan matplotlib
5. Plotly: visualisasi interaktif berbasis web

Anaconda Data Platform ini adalah kita dapat dengan mudah melakukan konfigurasi semua library yang berhubungan di satu tempat. Kelebihan lain, di Anaconda juga menyediakan akses ke tools data science, dan pengembangan software lainnya seperti Spyder, Microsoft Visual Studio Code, dan juga R Studio bagi anda yang juga terbiasa dengan bahasa pemrograman R (Wasistiono et al., 2020). upyter Notebook adalah sistem yang paling banyak digunakan untuk pemrograman interaktif yang dirancang untuk membuat analisis data lebih mudah untuk didokumentasikan, dibagikan, dan direproduksi. Jupyter berasal dari IPython dan selain Python, mendukung berbagai bahasa pemrograman, seperti bahasa R, JavaScript, dan C. Jupyter Notebook (file yang berekstensi ipynb) adalah dokumen yang dihasilkan oleh Jupyter Notebook App yang berisikan kode komputer dan rich text element seperti kode dan teks (Pimentel et al., 2021) .

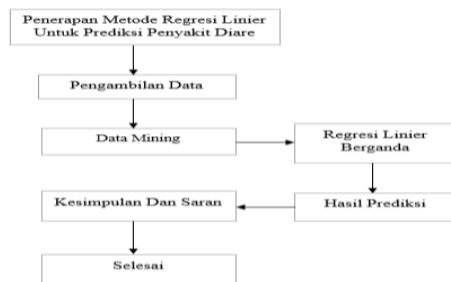
II. METODOLOGI PENELITIAN

Metode Penelitian yang digunakan adalah jenis penelitian deskriptif kuantitatif. Penelitian kuantitatif adalah suatu proses



menemukan pengetahuan yang menggunakan data berupa angka sebagai alat menganalisis keterangan mengenai apa yang ingin kita ketahui. penelitian kuantitatif sarat dengan nuansa angka-angka dalam teknik pengumpulan data di lapangan (Mujahidin & Pribadi, 2017).

Terdapat beberapa tahapan dalam pengerjaan, yaitu:



Gambar 1. Flowchart Penelitian

Metode Pengumpulan data dalam penelitian ini menggunakan data primer yaitu mengumpulkan data langsung dari objek yang diteliti, seperti pengamatan, wawancara, dan dokumentasi. Selain data primer juga menggunakan data skunder yaitu data yang didapat berasal dari berbagai sumber seperti buku, laporan, jurnal, dan lain-lain.

Metode Analisa yang di gunakan dalam penelitian ini adalah memgggunakan metode Regresi Linier Berganda dengan implementasi menggunakan bahasa Pemrograman Python. Dalam Regresi Linier, garis regresi dinyatakan dalam bentuk persamaan linier adapun rumus regresi linier berganda sebagai berikut (Syukron, 2017):

$$Y = a + b_1x_1 + b_2x_2 + \dots + b_nx_n \dots (1)$$

Dimana:

1. Y : Variabel dependen (nilai yang ingin kamu prediksi)
2. X, X₂,...X_n : Variabel independen (nilai yang digunakan untuk memprediksi)
3. a : Konstanta atau titik potong sumbu Y (intercept)

4. b₁, b₂,...b_n : Koefisien regresi yang sesuai dengan variabel independent.

Adapun metode pengujian yang dilakukan memprediksi, yaitu menggunakan *Mean Absolute Error* (MAE), *Mean Squere Error* (MSE), *Mean Precentage Absolute* (MAPE) (Bahtiar, 2023).

III. HASIL DAN PEMBAHASAN

1. Pengumpulan Data

Data penelitian yang digunakan adalah data jumlah masyarakat yang menderita penyakit diare dan variabel yang mempengaruhi yaitu suhu, udara dan kualitas air. Data penelitian yang digunakan adalah data data tahun 2019 sampai tahun 2023 .

No	Tahun	Suhu	Udara	Kualitas Air	Jumlah Pasien
1	Jan-19	26	29,2	42,73	23
2	Feb-19	28	38,3	31,27	25
3	Mar-19	31	25,1	32,99	29
4	Apr-19	28	24	32,21	30
5	May-19	26	24,2	22,68	21
6	Jun-19	24	29,6	31,27	24
7	Jul-19	23	22,5	34,71	26
8	Aug-19	24	26,6	38,14	28
9	Sep-19	26	34	32,99	30
10	Oct-19	28	16,3	32,21	29
11	Nov-19	31	17,5	34,71	35
12	Dec-19	28	20	42,27	31
13	Jan-20	26	17,3	42,73	26
14	Feb-20	24	20,3	31,27	24
15	Mar-20	23	20,1	37,37	19
16	Apr-20	24	19,2	34,71	23
17	May-20	27	22,8	22,68	27
18	Jun-20	29	21	31,27	26
19	Jul-20	32	19,4	34,71	21
20	Aug-20	29	17,9	38,14	29
21	Sep-20	27	21,1	24,4	30
22	Oct-20	25	32,1	34,71	27
23	Nov-20	23	47,1	42,27	29
24	Dec-20	24	19,9	42,73	25
25	Jan-21	24	29	42,73	27
26	Feb-21	31	34,1	31,27	24
27	Mar-21	25	25,4	32,99	20
28	Apr-21	23	25,6	32,21	22
29	May-21	24	26,9	32,99	20
30	Jun-21	30	25,5	32,21	26
31	Jul-21	25	24,8	34,71	25
32	Aug-21	23	26,4	42,27	21
33	Sep-21	24	14,2	42,73	23
34	Oct-21	30	11,9	31,27	20
35	Nov-21	27	14,3	37,37	29
36	Dec-21	24	14,5	34,71	22

32	Aug-21	23	26,4	42,27	21
33	Sep-21	24	14,2	42,73	23
34	Oct-21	30	11,9	31,27	20
35	Nov-21	27	14,3	37,37	29
36	Dec-21	24	14,5	34,71	22
37	Jan-22	24	13,9	22,68	26
38	Feb-22	29	16	31,27	19
39	Mar-22	26	10,5	34,71	23
40	Apr-22	24	13,8	38,14	21
41	May-22	26	13,6	24,4	20
42	Jun-22	29	20,3	42,27	25
43	Jul-22	27	20,1	42,73	22
44	Aug-22	25	19,2	31,27	29
45	Sep-22	23	22,8	37,37	24
46	Oct-22	24	21	34,71	30
47	Nov-22	24	19,4	22,68	26
48	Dec-22	31	17,9	31,27	28
49	Jan-23	25	21,1	34,71	21
50	Feb-23	23	32,1	38,14	20
51	Mar-23	24	47,1	24,4	26
52	Apr-23	30	19,9	34,71	24
53	May-23	26	13,8	22,68	27

Gambar 2. Data Set

2. Deklarasi Library Python

Library pada bahasa pemrograman python berfungsi agar kita dapat menggunakan fungsi dari library tersebut. Tampilan deklarasi library python dapat dilihat pada gambar berikut:

```
import numpy as np
import pandas as pd
import matplotlib as plt
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import MinMaxScaler, StandardScaler
from sklearn.impute import SimpleImputer
from sklearn import metrics
import seaborn as sns
import seaborn as sbn
import warnings
warnings.filterwarnings("ignore")
```

Gambar 3. Tampilan Library Python

3. Read_CSV

Read_csv pada bahasa pemrograman python digunakan untuk mengakses data set dalam format CSV (Comma Separated Values). Perintah read_excel dapat dilihat pada gambar berikut:

```
diare = pd.read_csv('Data_denci.csv')
diare.head()
```

Gambar 4. Perintah Read_CSV

4. Missing Value

Terdapat beberapa cara untuk menangani nilai yang hilang (missing value) dalam dataset. Beberapa fungsi atau perintah yang berguna untuk menangani nilai yang hilang (missing value) dapat dilihat pada gambar berikut:

```
# Mengecek apakah ada nilai yang hilang
print('\nMengecek Missing Values')
display(diare.isnull().sum())
```

Mengecek Missing Values

```
Suhu          0
Udara         0
Kualitas_Air  0
Jumlah_Pasien 0
dtype: int64
```

Gambar 5. Tampilan Missing Value

5. Model Persamaan Regresi Linier

Model persamaan regresi linier dibangun berdasarkan 2 variabel, yaitu variabel independen (suhu, udara dan kualitas air) dan variabel dependen (jumlah pasien). Perintah python yang digunakan untuk membuat model persamaan regresi linier sebagai berikut:

```
#model persamaan regresi berganda
reg = LinearRegression()
reg.fit(diare[['Suhu', 'Udara', 'Kualitas_Air']], diare.Jumlah_Pasien)
```

Gambar 6. Model Regresi Linier

```
#mencari nilai a
reg.intercept_

13.116333474572714

#mencari nilai b1,b2 dan b3
reg.coef_

array([ 0.39001012,  0.07766828, -0.001093  ])
```

Gambar 7. Hasil Persamaan Regresi Linier

6. Tes Persamaan Model Regresi dengan Data Uji

Model regresi yang terbentuk dapat kita ujikan dengan data baru seperti pada gambar berikut:


```
#tes prediksi
reg.predict([[23,40,11]])

array([25.18127441])
```

Gambar 8. Tampilan Tes Model Regresi

7. Hubungan Antara Variabel X dan Y (Scatter Plot)

Scatter plot adalah jenis grafik yang digunakan untuk menampilkan hubungan antara dua variabel numerik dalam bentuk titik-titik di dalam bidang koordinat. Scatter plot sangat berguna untuk menunjukkan pola, hubungan, atau korelasi antara dua variabel:

```
# Membuat scatter plot antara fitur dan target
fig, axes = plt.subplots(1, 3, figsize=(15, 5))
axes[0].scatter(diare['Suhu'], diare['Jumlah_Pasien'], color='blue')
axes[0].set_xlabel('Suhu')
axes[0].set_ylabel('Jumlah_Pasien')
axes[0].set_title('X1 vs y')
axes[1].scatter(diare['Udara'], diare['Jumlah_Pasien'], color='green')
axes[1].set_xlabel('Udara')
axes[1].set_ylabel('Jumlah_Pasien')
axes[1].set_title('X2 vs y')
axes[2].scatter(diare['Kualitas_Air'], diare['Jumlah_Pasien'], color='red')
axes[2].set_xlabel('Kualitas_Air')
axes[2].set_ylabel('Jumlah_Pasien')
axes[2].set_title('X3 vs y')
plt.tight_layout()
plt.show()
```

Gambar 9. Tampilan Perintah Scatter Plot

8. Koefisien Korelasi

Koefisien korelasi adalah metrik statistik yang digunakan untuk mengukur sejauh mana dua variabel terkait satu sama lain dalam suatu dataset. Koefisien korelasi mengukur kekuatan dan arah hubungan antara dua variabel.

```
# Menghitung korelasi antara fitur dan target
correlations = diare.corr()
print(correlations['Jumlah_Pasien'])

Suhu          0.249802
Udara          0.111739
Kualitas_Air  -0.026122
Jumlah_Pasien  1.000000
Name: Jumlah_Pasien, dtype: float64
```

Gambar 10. Tampilan Hasil Koefisien Korelasi

9. Grafik Hubungan Antar Variabel dengan Pairplot

Fungsi *pairplot* biasanya ada di library Seaborn. Ini digunakan untuk membuat grid dari sejumlah variabel numerik dalam bentuk scatter plot dan histogram. *pairplot* sangat berguna untuk mengeksplorasi hubungan antara berbagai variabel dalam dataset yang sama.

```
# Membuat data sintetis/ hubungan variabel independen dan variabel dependen
np.random.seed(0)
x1 = np.random.normal(0, 1, 100)
x2 = np.random.normal(0, 1, 100)
x3 = np.random.normal(0, 1, 100)
y = 3 + 2*x1 + 1.5*x2 + 0.5*x3 + np.random.normal(0, 0.5, 100)
diare = pd.DataFrame({'Suhu': x1, 'Udara': x2, 'Kualitas_Air': x3, 'Jumlah_Pasien': y})
sns.pairplot(diare)
plt.show()
```

Gambar 11. Tampilan Tes Model Regresi

10. Uji Model Regresi

Uji model dilakukan untuk melihat perbedaan antara data actual dengan data prediksi. Perintah uji model dapat dilihat pada gambar berikut:

```
#Uji Model
Reg=LinearRegression()
Reg.fit(X_test,y_test)

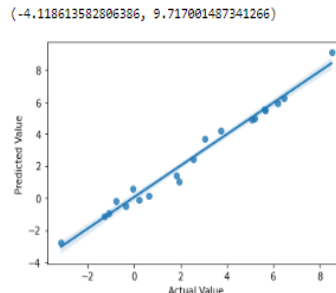
LinearRegression()
```

Gambar 12. Uji Model Regresi Linier Berganda

11. Grafik Model Regresi Linier Berganda

Grafik model regresi linier berganda menggunakan perintah `sns.regplot` digunakan untuk menggambarkan model regresi linier berganda yang terbentuk. Grafik model regresi linier berganda dapat dilihat pada gambar berikut:

```
#Visualisasi hasil data aktual dengan data prediksi
sns.regplot(x="Actual Value", y="Predicted Value", data=pred_df)
plt.ylim()
```



Gambar 13. Grafik Model Regresi Linier Berganda

12. Uji Model Regresi Linier Berganda

Pengujian model regresi linier yang dilakukan untuk melihat seberapa baik model yang terbentuk nantinya akan digunakan untuk memprediksi jumlah penderita penyakit diare dimasa yang akan datang. Pengujian model regresi linier dapat dilihat pada gambar berikut:

```
#Mencari Nilai Mae,MSE,RMSE dan MAPE
mae = metrics.mean_absolute_error(y_test, y_pred)
mse = metrics.mean_squared_error(y_test, y_pred)
r2 = np.sqrt(metrics.mean_squared_error(y_test, y_pred))

def mape(y_test, pred):
    y_test, pred = np.array(y_test), np.array(pred)
    mape = np.mean(np.abs((y_test - pred) / y_test))
    return mape
LR_Test_predict = reg.predict(X_train)
LR_MAPE= mape(y_train,LR_Test_predict)

print('Mean Absolute Error:', mae)
print('Mean Square Error:', mse)
print('r2 score:', r2)
print("MAPE: ", LR_MAPE)

Mean Absolute Error: 0.35609660506590435
Mean Square Error: 0.17408336723384799
r2 score: 0.4172329891485667
MAPE: 7.575155367200233
```

Gambar 14. Uji Model Regresi Berganda

IV. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan dan telah dijelaskan di atas maka penulis dapat menarik kesimpulan bahwa :

1. Adanya sistem prediksi jumlah pasien penyakit diare, maka UPT Puskesmas Terawas dapat memprediksi jumlah pasien penyakit diare di masa mendatang, sehingga dapat mempersiapkan segala kebutuhan seperti obat-obatan dan Tindakan lainnya untuk menangani penyakit diare di UPT Puskesmas Terawas.
2. Hasil pengujian didapat nilai *Mean Absolute Error* (MAE) sebesar 0.35609660506590435, Nilai *Mean Square Error* (MSE) sebesar 0.17408336723384799, Nilai R² sebesar 0.4172329891485667 dan MAPE 7.575155367200233. Dari hasil pengujian yang dilakukan dapat dikatakan bahwa model regresi linier yang dihasilkan berkategori “Sangat Baik”.

V. SARAN

Pengembangan sistem prediksi ini hendaknya ditambahkan variabel prediktor lainnya yang berkaitan dengan faktor-faktor yang berkaitan dengan penyakit diare sehingga mendapat hasil yang lebih akurat.

VI. DAFTAR PUSTAKA

- Bahtiar, A. (2023). Jurnal Informatika Terpadu PREDIKSI HASIL PANEN PADI TAHUN 2023 MENGGUNAKAN METODE REGRESI LINIER DI KABUPATEN INDRAMAYU. *Jurnal Informatika Terpadu*, 9(1), 18–23. <https://journal.nurulfikri.ac.id/index.php/JIT>
- Data, P., Kementerian, I., Republik, K., & Kunci, K. (2023). *PENERAPAN MACHINE LEARNING DALAM PREDIKSI TINGKAT KASUS PENYAKIT DI INDONESIA* Master of Information Systems Management Bina Nusantara University Abstraksi Keywords: *Pendahuluan Tinjauan Pustaka Metode Penelitian Proses implementasi machine learning dapat.* 5(1).
- Gede, D., Apriani, Y., Made, D., Sastra, F., & Sri, N. (2022). 3 1,2,3. 1, 15–26.
- Hermawan, A. K., & Nugroho, A. (2023). Analisa Data Mining Untuk Prediksi Penyakit Ginjal Kronik Dengan Algoritma Regresi Linier. *Bulletin of Information Technology (BIT)*, 4(1), 37–48.
- Ibrahim, I., & Sartika, R. A. D. (2021). Faktor-Faktor yang Berhubungan dengan Kejadian Diare pada Siswa Sekolah Dasar di Kabupaten Lebak, Provinsi Banten, Indonesia. *Indonesian Journal of Public Health Nutrition*, 2(1), 34–43. <https://doi.org/10.7454/ijphn.v2i1.5338>



- Kafil, M., & Industri, F. T. (2019). *PENERAPAN METODE K-NEAREST NEIGHBORS*. 3(2), 59–66.
- Linier, K., Data, T., Regresi, U., Principal, M., & Analysis, C. (2023). *Jurnal Teknologi Terpadu*. 9(1), 1–9.
- Mujahidin, A., & Pribadi, D. (2017). Penerapan Algoritma C4.5 Untuk Diagnosa Penyakit Pneumonia Pada Anak Balita Berbasis Mobile. *Jurnal Swabumi*, 5(2), 155–161.
- Normah, Rifai, B., Vambudi, S., & Maulana, R. (2022). Analisa Sentimen Perkembangan Vtuber Dengan Metode Support Vector Machine Berbasis SMOTE. *Jurnal Teknik Komputer AMIK BSI*, 8(2), 174–180. <https://doi.org/10.31294/jtk.v4i2>
- Penyakit, A., Melitus, D., & Kota, D. I. (2019). *THE APPLICATION OF BACKPROPAGATION TO PREDICT*. 1, 51–55.
- Pimentel, J. F., Murta, L., Braganholo, V., & Freire, J. (2021). Understanding and improving the quality and reproducibility of Jupyter notebooks. *Empirical Software Engineering*, 26(4). <https://doi.org/10.1007/s10664-021-09961-9>
- Santos, T. B. (2019). Aplikasi Data Mining untuk Clustering Daerah Penyebaran Penyakit Diare di DKI Jakarta Menggunakan Algoritma K-MEANS. *Jurnal Ilmiah FIFO*, 11(2), 131. <https://doi.org/10.22441/fifo.2019.v11i2.003>
- Shidiq, A. R., Ermawati, I., & Santoni, M. M. (2020). *Penerapan Algoritma Regresi Linier Untuk Prediksi Pada Data Obat Antibiotik*. 102–111.
- Simbolon, C. E. (2021). Penerapan Algoritma Regresi Linier Sederhana Dalam Memprediksi Keuntungan dan Kerugian Kelapa Sawit Pt . Sri Ulina. *Journal of Information System Research (JOSH)*, 2(2), 169–172.
- Syafruddin, M., Hakim, L., & Despa, D. (n.d.). *Metode Regresi Linier untuk Prediksi KebutuhanEnergi Listrik Jangka Panjang (Studi Kasus Provinsi Lampung). 1*.
- Sylfi, dwi ispriyanti, diah safitri. (2017). Analsis Regresi. *Jurnal Gaussin*, 1.
- Syukron, B. (2017). Kata kunci . *Kinabalu*, 11(2), 50–57.
- Wahid, M. A. R., Nugroho, A., Anshor, A. H., Teknik, F., Informatika, T., & Bangsa, P. (2023). *Prediksi Penyakit Kanker Paru-Paru Dengan Algoritma Regresi Linier*. 4(1), 63–74.
- Wasistiono, S., Indrayani, E., & Pitono, A. (2020). Document Title/. *Quality, March*, 1–6.
- Wijayadhi, A., Effendi, M. M., & Rahardjo, S. B. (2023). *Prediksi Penyakit Jantung Dengan Algoritma Regresi Linier*. 4(1), 15–28.
- Yuanita, A., & Yuanita, A. (2023). *Penerapan Metode Regresi Linier Berganda Pada Kasus Balita Gizi Buruk Di Kabupaten Bojonegoro*.