

Prediksi Jumlah Produksi Sablon Tahun Menggunakan Algoritma Regresi Linear di Nolbas SVN

Muhammad Fadhilah¹, Martanto², Irfan Ali³

¹Teknik Informatika, STMIK IKMI Cirebon, Indonesia

²Manajemen Informatika, STMIK IKMI Cirebon, Indonesia

³Rekayasa Perangkat Lunak, STMIK IKMI Cirebon, Indonesia

fadhilahmuhammad740@gmail.com

Info Artikel

Sejarah artikel:

Diterima Mei 2023

Direvisi Juni 2023

Disetujui Juli 2023

Diterbitkan Juli 2023

ABSTRACT

Nolbas svnr is a business engaged in the clothing industry which refers more to t-shirt screen printing. This business carries out its activities based on customer orders received through orders from individuals, shops, and schools. With the many types of screen printing that are made, the number of orders received and executed by Nolbas Svnr increases. Screen printing production at Nolbas Svnr is always changing every year. The main objective of this research is to obtain a predictive model for the amount of screen printing production using the Linear Regression method based on the number of orders obtained each year. The results that can be obtained in research can help for the supply of raw materials, the amount of raw materials, paint and so on. This study uses the linear regression method to process sales data using attributes such as year, customer name, price of goods, price of materials and the number of orders. of 0.5601. The results of the constant values and regression coefficients are used to predict the amount of screen printing production in 2023 at Zerobas SVN and the predicted value is 3391. Evaluation of the multiple linear regression model shows an MAE value of 3.7247, an MSE value of 17.8633 and an R2 score of 87% .

Keywords: Linear Regression; Screen Printing Production; Nolbas SVN.

ABSTRAK

Nolbas SVN merupakan usaha yang bergerak di bidang industri pakaian yang lebih mengacu pada sablon kaos. Usaha ini melakukan kegiatannya berdasarkan pesanan pelanggan yang diterima melalui order dari perorangan, toko, maupun sekolahan. Dengan banyaknya type sablon yang dibuat semakin banyak pula jumlah orderan yang diterima dan dikerjakan oleh nolbas svnr. Produksi sablon di Nolbas SVN tiap tahun senantiasa berubah- ubah. Tujuan utama dari penelitian ini adalah untuk memperoleh model prediksi jumlah produksi sablon dengan menggunakan metode Regresi Linear berdasarkan banyak orderan diperoleh tiap tahunnya. Hasil yang bisa didapatkan pada penelitian dapat membantu untuk persediaan bahan baku, jumlah bahan baku, cat dan sebagainya. Penelitian ini menggunakan metode regresi linear untuk mengolah data penjualan menggunakan atribut seperti tahun, nama pelanggan, harga barang, harga bahan dan jumlah pesanan. Hasil pengujian didapatkan konstanta regresi nilai positif sebesar 179.4432, nilai koefisien B1 adalah -2.4839 dan nilai koefisien B2 adalah sebesar 0.5601. Hasil nilai konstanta dan koefisien regresi tersebut digunakan untuk melakukan prediksi jumlah produksi sablon tahun 2023 di Nolbas SVN dan nilai prediksinya adalah 3391. Evaluasi dari model regresi linear berganda di dapatkan nilai MAE sebesar 3.7247, nilai MSE sebesar 17.8633 dan nilai r2 score didapatkan sebesar 87%.

Kata Kunci : Produksi Sablon; Regresi Linear; Nolbas SVN.

PENDAHULUAN

Dalam dunia bisnis semuanya bergerak sangat cepat baik dalam skala kecil maupun skala besar begitupun dalam sektor industri seperti UMKM. UMKM merupakan usaha yang berperan besar dalam membantu perekonomian Indonesia tetap bertahan, mereka diandalkan untuk menyediakan banyak layanan dan produk yang dibutuhkan. Salah satu bisnis yang berkembang pesat saat ini adalah sablon karena ada banyak permintaan besar untuk produk yang seragam dan bermanfaat [1]. Nobas svnr adalah usaha konveksi sablon kaos di Cirebon yang menyediakan jasa sablon dan sablon kaos sesuai pesanan pelanggan. Rata-rata perusahaan memproduksi 100-150 kaos per minggu. Jumlah ini pasti akan bertambah seiring pertumbuhan perusahaan. Seiring perkembangannya, perusahaan sablon kaos ini mulai merasakan adanya masalah seperti proses pemesanan, produksi, dan pengumpulan data keuangan dalam banyak [2].

Produksi sablon di nolbas svnr tiap tahun senantiasa berubah-ubah. Pada tahun 2020 produksi sablon diketahui sebanyak 2.975pcs. serta pada tahun 2021 diketahui produksi sablon menurun sebesar 2.643pcs. Sebaliknya pada tahun 2022 produksi sablon hadapi peningkatan ialah 3.550pcs. Terdapatnya peningkatan serta penyusutan ini diakibatkan oleh sebagian aspek, semacam kenaikan bahan baku, serta harga cat. Dalam riset ini penulis memakai dataset yang berasal dari penjualan nolbas svnr.

Berdasarkan penelitian terdahulu oleh D Novianti, N Dias Palasaran, M Qomaruddin, pada tahun 2017 yang berjudul “Algoritma Regresi Linear pada Prediksi Permohonan Paten yang Terdaftar di Indonesia”. Dari penelitian ini di dapatkan bahwa metode linear regresi layak dan efektif untuk memprediksi pendaftaran paten untuk tahun selanjutnya berdasarkan data pendaftaran dari tahun 2014 sampai tahun 2018. Berdasarkan hasil prediksi ditarik kesimpulan hasil prediksi untuk Taiwan pendaftaran paten di tahun selanjutnya sebesar 378 paten, Canada sebesar 1.031 paten, dan China sebesar 3.150 paten [3].

Perbandingan penelitian terdahulu dengan penelitian yang akan di teliti adalah tentang dataset nya dataset yang digunakan pada penelitian terdahulu menggunakan data Hak Kekayaan Intelektual dan untuk penelitin yang akan di teliti menggunakan data laporan penjualan sablon. Keunggulan penelitian yang penulis lakukan yaitu penerapan metode regresi linier pada prediksi jumlah hasil sablon di nolbas svnr, dimana penelitian ini belum pernah dilakukan di nolbas svnr.

Tujuan utama dari penelitian ini adalah untuk memperoleh model prediksi jumlah produksi sablon dengan menggunakan metode Regresi Linear berdasarkan banyak orderan diperoleh tiap tahunnya. Prediksi ini bermanfaat untuk membantu perusahaan dalam memprediksi jumlah produksi tiap tahunnya. Hasil yang bisa didapatkan pada penelitian dapat membantu untuk persediaan bahan baku, jumlah bahan baku, cat dan sebagainya.

METODE

Metode yang digunakan pada penelitian ini yaitu menggunakan metode regresi linier. Regresi Linier adalah algoritma yang digunakan untuk mengukur hubungan antara korelasi dua variabel atau lebih yang digunakan untuk prediksi

melalui garis lurus. Variabel sendiri merupakan ukuran yang memiliki nilai yang berubah ubah[5]. Model persamaan regresi linier berganda adalah sebagai berikut:

$$Y=a+b_1x_1+b_2x_2$$

Dimana :

Y = Variabel terikat

a = nilai intercept/konstanta

b = koefisien regresi

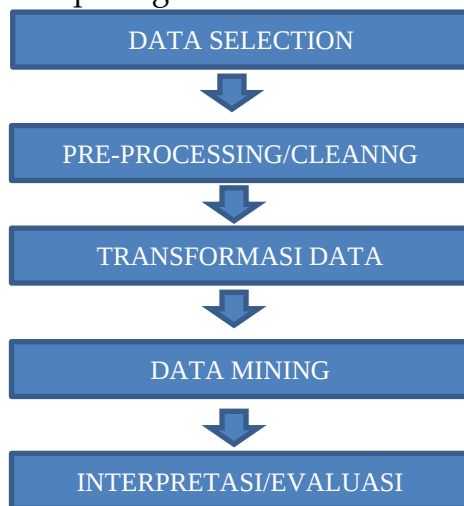
X = variabel bebas

Berikut ini rumus yang digunakan untuk menentukan nilai a dan b :

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

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

Pada tahapan penelitian dilakukan tahapan-tahapan khusus dalam data mining yang lazim disebut dengan *Knowledge Data Discovery*, *Knowledge Data Discover* merupakan proses analisa terstruktur untuk memperoleh informasi yang mendekati kebenaran, baru, bermanfaat, dan menemukan pola dari data yang besar dan komplek. Data *mining* (DM) menjadi inti dari proses KDD, yaitu dengan menggunakan algoritma tertentu untuk mengeksplorasi data, membangun model dan menemukan pola yang belum diketahui. Model digunakan untuk memahami fenomena data, analisa maupun prediksi [6]. langkah/tahap dari proses *Knowledge Data Discovery* dapat dilihat pada gambar1.



Gambar 1. *Knowledge Data Discovery*

HASIL DAN PEMBAHASAN

Data Selection

Data yang digunakan dalam penelitian ini yaitu data produksi selama 3 tahun terakhir dari 2020-2022 yang berasal dari nolbas svnr kemudian data tersebut diseleksi dan akan digunakan untuk diolah dalam memprediksi jumlah produksi pada tahun 2023. Adapun atribut yang digunakan dalam penentuan prediksi jumlah produksi adalah atribut tahun, nama_konsumen, jumlah, harga_barang, harga_bahan. Dapat dilihat pada Tabel 1.

Tabel 1. Data Selection

No	Tahun	Nama_Konsumen	Jumlah	Harga_Barang	Harga_Bahan
0	2020	Pejuang Sodaqoh	3	95	93
1	2020	Hj.Sarah	18	87	88
2	2020	Andi (Tuk)	25	80	80
3	2020	Bohel (Kalitengah)	31	75	72
4	2020	Diah/Ibu (Sumber)	60	50	43
5	2020	Erum	63	50	43
6	2020	Fajar (Cirebon)	63	48	43
7	2020	Fajar (Cirebon)	64	47	41
8	2020	Indocement	75	40	32
9	2020	Ikkal	17	88	93
10	2020	H.sarah	18	87	88
.....					
178	2022	Pak Dosen	100	35	28

Pre-processing/Cleaning

Tahap preprocessing merupakan tahap awal dari proses KDD. Pada tahapan ini data yang missing value harus dibersihkan. Hal ini dikarenakan data yang missing value merupakan syarat awal dalam melakukan data mining. Suatu data dikatakan missing value jika terdapat atribut dalam dataset yang tidak berisi nilai atau kosong. Dapat dilihat pada Gambar 2 dan Gambar 3.

```
In [6]: sablondf.isnull().sum()
```

```
Out[6]: tahun          0
        nama_konsumen  0
        jumlah        0
        harga_barang  0
        harga_bahan    0
        dtype: int64
```

Gambar 2. Data Missing Value

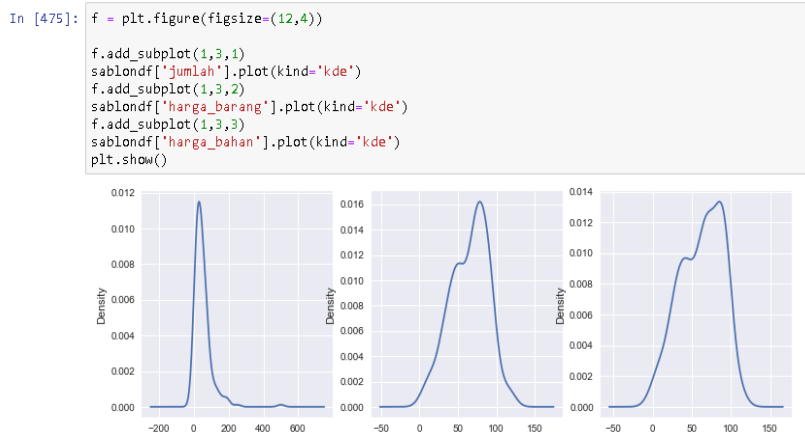
Transformation

Pada tahap ini tidak diperlukan transformasi karena *value* pada atribut yang digunakan semuanya telah berupa data numeric.

```
In [5]: sablondf.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 503 entries, 0 to 502
Data columns (total 5 columns):
tahun          503 non-null int64
nama_konsumen  503 non-null object
jumlah         503 non-null int64
harga_barang   503 non-null int64
harga_bahan    503 non-null int64
dtypes: int64(4), object(1)
memory usage: 19.7+ KB
```

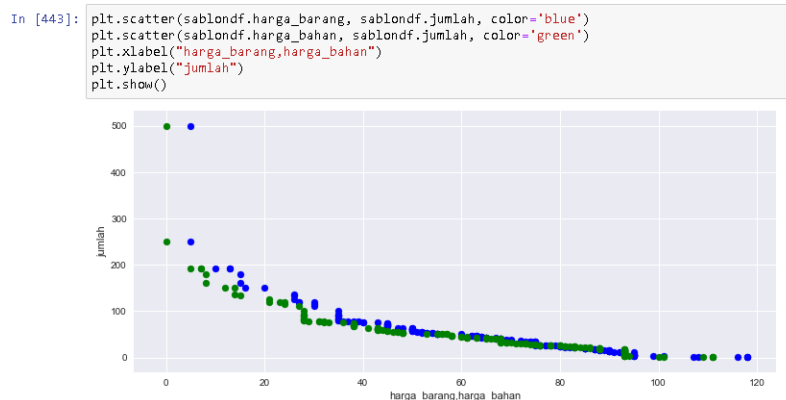
Gambar 3. Data Missing Value



Gambar 1. Hubungan Variabel X dan Variabel Y

Data Mining

Berikut data mining tentang hubungan variabel X dan variabel Y.



Gambar 5. Hubungan Variabel X dan Variabel Y

a. Membuat Model Regresi Linear

Untuk membuat model regresi dalam *python Library* yang digunakan adalah *numpy*, *pandas*, *seaborn*, *matplotlib*. Sedangkan model *stuff* yang digunakan dari *sklearn.model_selection import train_test_split*, *sklearn*. *Linear_model import LinearRegression*, *sklearn import metrics* dapat dilihat pada Gambar 6.

```
In [2]: import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# Model stuff #
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn import metrics
import statsmodels.api as sm

import os
print(os.listdir("C:/Users/ASUS/Documents/Zoom/"))

import warnings
warnings.filterwarnings("ignore")

%matplotlib inline
plt.style.use("seaborn")
plt.rcParams['figure.figsize'] = (12,5)
```

Gambar 6. Import Library

Berikut merupakan penjelasan library pada Gambar 6:

- 1) *Pandas* digunakan untuk membaca data dan menampilkan data.
 - 2) *Numpy* digunakan untuk mendukung operasi *array*.
 - 3) *Matplotlib* digunakan untuk visualisasi data.
 - 4) *Sklearn* digunakan untuk proses prediksi. *Library Sklearn* mempunyai banyak *package* yang digunakan dalam proses prediksi menggunakan regresi linear sederhana
 - 5) *Sklearn.linear_model import LinearRegression* digunakan untuk memanggil algoritma *LinearRegression*.
 - 6) *Sklearn.model_selection import train_test_split* digunakan untuk membagi data menjadi training dan testing.
- b. Menentukan Variabel X dan variabel Y
- Membedakan data yang jadi variabel dependen serta variabel independen. Kolom yang jadi variabel y merupakan jumlah serta variabel x merupakan harga_barang, harga_bahan.

```
In [438]: feature_names=['harga_barang','harga_bahan']
X=sablondf[feature_names]
X
y=sablondf.jumlah
```

Gambar 7. Proses Menentukan Variabel X dan Variabel Y

- c. Proses Data Test dan Data Training

Pada tahapan ini dilakukan proses input data test dan data training yang dibagi menjadi data test 20% dan data *training* 80%. Dapat dilihat pada Gambar 8.

```
In [444]: X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=0)

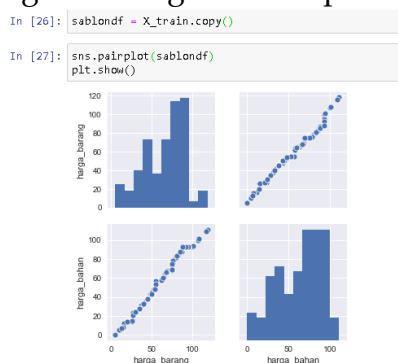
In [445]: print(X_train.shape)
print(y_train.shape)
print(X_test.shape)
print(y_test.shape)

(143, 2)
(143,)
(36, 2)
(36,)
```

Gambar 8. Pembagian Data Test dan Data Training

- d. Data Analisis

Analisis data eksplorasi diperlukan sebelum membangun model prediksi dan memungkinkan data implisit pola, yang membantu dalam memilih pembelajaran mesin yang sesuai algoritma dapat dilihat pada Gambar 9.



Gambar 9. Eksperimen untuk Menemukan Kombinasi Terbaik

e. Korelasi Antar Data

Analisa korelasi digunakan untuk melihat keterkaitan tentang derajat ikatan variabel sehingga bisa mengenali ikatan variabel yang terdapat. Bisa dilihat pada Gambar 10.

```
In [442]: sablondf.corr().style.background_gradient().set_precision(1)
```

```
Out[442]:
```

	tahun	jumlah	harga_barang	harga_bahan
tahun	1	0.08	-0.02	-0.03
jumlah	0.08	1	-0.8	-0.8
harga_barang	-0.02	-0.8	1	1
harga_bahan	-0.03	-0.8	1	1

Gambar 10. Perintah dan Hasil Korelasi Data

f. Rumus Regresi Linear

Setelah melakukan pembagian data test dan data training proses selanjutnya adalah memasukan rumus regresi linear. Dapat dilihat pada Gambar 11.

```
In [447]: regressor = LinearRegression()
regressor.fit(X_train, y_train)
```

```
Out[447]: LinearRegression(copy_X=True, fit_intercept=True, n_jobs=None, normalize=False)
```

Gambar 11. Proses Rumus Regresi Linear

Langkah selanjutnya adalah menghitung regresi linear . Dapat dilihat pada Gambar 11 dan Gambar 12.

```
In [450]: print(regressor.intercept_)
```

```
179.44321193406117
```

```
In [449]: print(regressor.coef_)
```

```
[-2.48394925  0.5601043 ]
```

Gambar 2. Proses Perhitungan Nilai Koefisien

Berdasarkan hasil pada persamaan regresi linear yaitu :

$$Y = 179.4432 + (-2.4839 * \text{harga_barang}) + (0.5601 * \text{harga_bahan})$$

g. Hasil prediksi metode regresi linear

Pada sesi ini hendak dicoba perbandingan antara nilai y awal dengan nilai y prediksi. Nilai y prediksi dihitung pada rumus regresi linear berganda yang sudah dicoba.

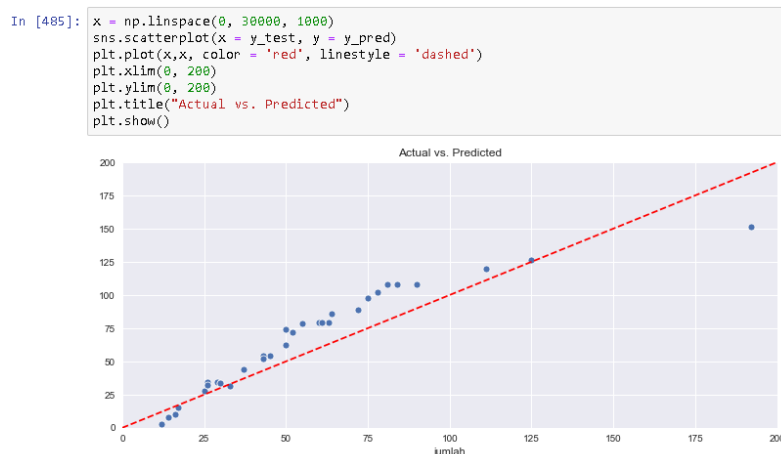
```
In [448]: y_pred=regressor.predict(X_test)
          dataframe = pd.DataFrame({'data sebenarnya': y_test, 'data prediksi' : y_pred})
          dataframe
```

```
Out[448]:
```

	data sebenarnya	data prediksi
137	33	31.234110
7	64	85.661873
124	29	34.594736
71	37	43.653857
135	50	74.191844
169	33	31.234110
56	25	28.019565
143	12	3.009581
161	26	31.867255
107	72	88.949459
33	45	54.636822

Gambar 3. Hasil Perhitungan Membandingkan Data Asli dengan Data Prediksi

Hasil proses pada Gambar 14 dapat dibuat dalam bentuk grafik. Hasil grafik memperlihatkan antara data asli dan data prediksi.



Gambar 4. Grafik Y Actual vs Y Prediksi

Evaluasi

Evaluasi untuk model regresi linear dapat dilakukan dengan beberapa cara evaluasi seperti MAE, MSE dan r2. Gambar 15. merupakan hasil perhitungannya.

```
In [452]: print(np.sqrt(metrics.mean_squared_error(y_test,y_pred)))
```

17.863394096075766

```
In [453]: print(np.sqrt(metrics.mean_absolute_error(y_test,y_pred)))
```

3.7247346913287593

```
In [451]: print(np.sqrt(metrics.r2_score(y_test,y_pred)))
```

0.874332061261199

Gambar 15. Hasil Perhitungan MSE, MAE, r2

Berdasarkan hasil pengujian didapatkan konstanta regresi nilai positif sebesar 179.4432, tanda positif artinya menunjukkan pengaruh yang searah antara variabel independen dan variabel dependen, hal ini menunjukkan bahwa jika semua variabel independen yang meliputi harga_barang, harga_bahan bernilai 0 atau tidak mengalami perubahan, maka nilai produksi adalah 179.4432.

Adapun nilai koefisien B1 adalah -2.4839 yakni nilai koefisien untuk variabel harga_barang. Nilai ini menunjukkan bahwa harga_barang berbanding terbalik dengan nilai produksi. Artinya jika variabel harga_barang mengalami kenaikan 1 satuan maka sebaliknya nilai produksi mengalami penurunan sebesar 2.4839. Nilai koefisien B2 adalah sebesar 0.5601 yakni nilai koefisien untuk variabel harga_bahan. Berbeda dengan harga_barang, harga_bahan memiliki nilai yang sejajar dengan nilai produksi. Dimana jika variabel harga_bahan mengalami kenaikan 1 satuan maka nilai produksi mengalami kenaikan sebesar 0.5601. Sehingga rumus regresi yang dihasilkan adalah dan dapat digunakan untuk prediksi selanjutnya adalah :

$$Y = 179.4432 + (-2.4839 * \text{harga_barang}) + (0.5601 * \text{harga_bahan})$$

Pada bagian ini melakukan analisis terhadap hasil prediksi dari proses pengujian yang telah dilakukan menggunakan Regresi Linear serta mengetahui hasil *mean absolute error* (MAE), *mean squared error* (MSE) dan r^2 untuk pengukuran performa dari hasil prediksi jumlah produksi di nolbas svnr.

```
In [453]: print(np.sqrt(metrics.mean_absolute_error(y_test,y_pred)))
3.7247346913287593
```

Gambar 5. Hasil Perhitungan MAE

Pada Gambar 16 ditunjukkan hasil MAE yang didapat dari hasil proses pengujian dimana nilai terendah menggunakan MAE sebesar 3.7247. Nilai MAE yang semakin rendah mengindikasikan bahwa nilai kesalahan semakin kecil yang berarti model bisa memprediksi dengan baik karena model dapat melakukan prediksi dengan selisih antara nilai aktual dan nilai prediksi yang tidak terlalu besar.

```
In [452]: print(np.sqrt(metrics.mean_squared_error(y_test,y_pred)))
17.863394096075766
```

Gambar 6. Hasil Perhitungan MSE

Pada Gambar 17 Ditunjukan hasil MSE yang didapat dari hasil proses pengujian dimana nilai terendah menggunakan MSE 17.8633. Metode Mean Squared Error secara umum digunakan untuk mengecek estimasi berapa nilai kesalahan pada peramalan. Nilai Mean Squared Error yang rendah atau nilai mean squared error mendekati nol menunjukkan bahwa hasil peramalan sesuai dengan data aktual dan bisa dijadikan untuk perhitungan peramalan di periode mendatang.

```
In [451]: print(np.sqrt(metrics.r2_score(y_test,y_pred)))
0.874332061261199
```

Gambar 7. Hasil Perhitungan r2

Pada Gambar 18 Menunjukan hasil nilai r2 yang didapat dari pengujian yang dimana nilai nya adalah 87%. Semakin tinggi nilai R2, semakin baik model prediksi dari model penelitian yang diajukan. Uji koefisien determinasi (R2) dilakukan untuk mengetahui dan memprediksi seberapa besar atau signifikan pengaruh gabungan variabel independen terhadap variabel dependen. Nilai koefisien determinasi bervariasi dari 0 sampai 1. Nilai yang mendekati 1 berarti variabel bebas menyediakan hampir semua informasi yang diperlukan untuk memprediksi variabel terikat. Namun jika nilai R2 menurun, berarti kemampuan variabel independen untuk menjelaskan variabel dependen sangat terbatas.

```
In [23]: sablondf = pd.DataFrame({
        "Data Prediksi": [54,73,88,167,31,52,-51,34,31,0,48,108,108,31,17,74,167,55,74,157,31,46,76,31,36,108,3,79,100,-51,12,3]
    })

In [24]: total = sablondf["Data Prediksi"].sum()
print(total)
3391
```

Gambar 8. Hasil Jumlah Produksi Tahun 2023

Pada Gambar 19 menunjukan dari proses prediksi yang telah dilakukan menghasilkan nilai jumlah produksi pada tahun 2023 di nolbas svnr yakni 3391.

PENUTUP

Berdasarkan hasil dan analisis diatas terkait prediksi jumlah produksi sablon 2023 menggunakan metode regresi linear berganda, maka didapatkan kesimpulan bahwa hasil dari regresi linear berganda untuk memprediksi jumlah produksi sablon tahun 2023 di nolbas svnr adalah 3391, serta hasil dari analisa regresi linear berganda di dapatkan nilai MAE sebesar 3.7247, nilai MSE sebesar 17.8633 dan nilai r2 score didapatkan sebesar 87%.

DAFTAR PUSTAKA

- [1] M. N. Rumpa, "Analisis Perhitungan Harga Pokok Produksi Pada Percetakan Maharani Multikreasi," pp. 1-5, 2020.
- [2] M. R. Abhad, A. Arwan, and D. Pramono, "Pengembangan Sistem Manajemen Perusahaan Sablon Kaos Berbasis Website Menggunakan Metode Prototyping (Studi Kasus: Perusahaan Sablon di Kota Malang)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 9, pp. 8514-8522, 2019.
- [3] D. Novianty, N. D. Palasara, and M. Qomaruddin, "Algoritma Regresi Linear pada Prediksi Permohonan Paten yang Terdaftar di Indonesia," *J. Sist. dan Teknol. Inf.*, vol. 9, no. 2, p. 81, 2021, doi: 10.26418/justin.v9i2.43664.
- [4] M. Masruroh, "Perbandingan Metode Regresi Linear Dan Neural Network Backpropagation Dalam Prediksi Nilai Ujian Nasional Siswa Smp

- Menggunakan Software R,” *Joutica*, vol. 5, no. 1, p. 331, 2020, doi: 10.30736/jti.v5i1.347.
- [5] V. P. Ramadhan and F. Y. Pamuji, “Analisis Perbandingan Algoritma Forecasting dalam Prediksi Harga Saham LQ45 PT Bank Mandiri Sekuritas (BMRI),” *J. Teknol. dan Manaj. Inform.*, vol. 8, no. 1, pp. 39–45, 2022, doi: 10.26905/jtmi.v8i1.6092.
- [6] Y. G. Siadari and D. Saripuna, “Data Mining Untuk Mengestimasi Jumlah Penumpang Pada Pt . Pinem Lau Guna Medan Dengan Menggunakan Metodere,” no. x, 2020.