

PREDIKSI JUMLAH PENGUNJUNG WISATAWAN BALI MENGUNAKAN DATA SET PENGUNJUNG DENGAN METODE LINEAR REGRESSION

Refiando Vandana¹, Mohammad Indra Fachrulrozy², Kevin Yong³

Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Tarumanagara

Jl. Letjen S. Parman No.1, Jakarta Barat, DKI Jakarta, Indonesia 11410

E-mail: ¹refiando.535180098@stu.untar.ac.id, ²mohammad.535180081@stu.untar.ac.id, ³

Kevin.535180107@stu.untar.ac.id

ABSTRAK

Pariwisata merupakan hal yang tidak bisa lepas dari kehidupan di Negara Indonesia. Berbagai tempat pariwisata tersebar di seluruh pelosok Indonesia. Salah satunya pariwisata Bali mempunyai manfaat salah satunya ekonomi. Manfaat pariwisata dari segi ekonomi di Indonesia adalah pariwisata dapat menghasilkan devisa yang besar bagi Negara sehingga meningkatkan perekonomian negara. Prediksi terhadap jumlah pengunjung wisatawan bali dengan menggunakan metode linear regression dapat membantu memprediksi pengunjung wisatawan bali, juga dapat membantu melihat peluang dalam sektor wisata dengan memberi gambaran-gambaran pengunjung wisatawan yang akan datang. Algoritma linear regresi dipilih karena merupakan algoritma peramalan yang mudah untuk diterapkan, tidak memerlukan desain eksperimen, atau teknik analisis yang canggih. Data diperoleh dari situs web resmi bali berbasis data pengunjung wisatawan bali.

Kata Kunci: Pemberlajaran Mesin, *Linear Regression*, Prediksi, Wisatawan, Bali

ABSTRACT

Tourism is an integral part of life in Indonesia. Various tourist attractions are scattered throughout Indonesia. One of them is Bali tourism, which has economic benefits. The economic benefits of tourism in Indonesia are that tourism can generate large foreign exchange earnings for the country, thereby boosting the country's economy. Predicting the number of tourists visiting Bali using the linear regression method can help predict the number of tourists visiting Bali, as well as help identify opportunities in the tourism sector by providing an overview of the tourists who will be visiting. The linear regression algorithm was chosen because it is an easy-to-apply forecasting algorithm that does not require experimental design or sophisticated analysis techniques. The data was obtained from the official Bali website based on data on tourists visiting Bali.

Keywords: Machine Learning, *Linear Regression*, Prediction, Tourist, Bali

1. PENDAHULUAN

1.1. Latar Belakang

Pariwisata merupakan hal yang tidak bisa lepas dari kehidupan di Negara Indonesia. Berbagai tempat pariwisata tersebar di seluruh pelosok Indonesia. Salah satunya pariwisata Bali mempunyai manfaat salah satunya ekonomi. Manfaat pariwisata dari segi ekonomi di Indonesia adalah pariwisata dapat menghasilkan devisa yang besar bagi Negara sehingga meningkatkan perekonomian negara. Devisa yang diterima secara berturut-turut pada tahun 1996, 1997, 1998, 1999, dan 2000 adalah sebesar 6,307.69; 5,321.46; 4,331.09; 4,710.22; dan 5,748.80 juta dollar AS (Santosa, 2001). Pada tahun 2002 dan 2003, meskipun mengalami tragedi Kuta (Bom Bali), nilai devisa juga masih tetap tinggi, yaitu US\$ 4.496 Milyard tahun 2002 dan US\$ 4.307 Milyard tahun 2003. Dari jumlah tersebut dapat dilihat bahwa pariwisata merupakan salah satu sumber devisa terbesar di Indonesia.

Padatny arus wisata di Bali harus diimbangi dengan fasilitas tempat wisata. Lonjakan jumlah wisatawan mancanegara (wisman) di Bali jika tidak diimbangi dengan fasilitas yang memadai

tentunya akan membuat para wisatawan kecewa dan mungkin tidak akan datang lagi ke Bali. Untuk itu dikembangkanlah suatu alat prediksi dimana alat tersebut akan mampu meramalkan jumlah kunjungan wisatawan mancanegara yang akan datang di Bali per tahun berdasarkan Negara asal dan per bulan berdasarkan pintu masuk dengan berbasiskan pada metode linear regression. Dengan mampu meramalkan jumlah kunjungan tersebut, diharapkan pemerintah Indonesia akan mampu mengoptimalkan persiapan dalam melakukan penyambutan wisatawan mulai dari segi fasilitas hingga keamanan.

1.2. Tujuan dan Manfaat

Penelitian ini bertujuan untuk :

- a. Memprediksi jumlah pengunjung yang akan datang pada bulan tertentu.
- b. Membantu dalam memprediksi jumlah pengunjung yang akan datang dari data set wisatawan bali.
- c. Membantu dalam melihat peluang pariwisata di bali dengan prediksi jumlah pengunjung wisatawan yang akan datang.
- d. Memberikan gambaran-gambaran mengenai prediksi jumlah pengunjung wisatawan bali yang akan datang.

1.3. Studi Literatur

1. Rancangan aplikasi pertama, Institut Teknologi Bandung pada studi kasus Fakultas Teknik Informatika dengan judul “*Predicting Football Match Results with Logistic Regression*”, oleh Darwin Prasetyo dan Dra. Harlili, M.Sc, penelitian ini mencoba melakukan prediksi hasil pertandingan sepak bola dengan menggunakan metode Logistic Regression yang memiliki prediksi kecocokan baik.
2. Rancangan aplikasi lanjut, *Department of Computer Science, Landmark University Omu Aran, Nigeria* dengan judul “*Predictive modelling of COVID-19 confirmed cases in Nigeria*”, oleh Roseline O. Ogundokun, Adewale F. Lukman, Golam B.M. Kibria, Joseph B. Awotunde, Benedita B. Aladeitan, penelitian ini mencoba melakukan prediksi covid-19 yang telah terkonfirmasi dengan menggunakan metode linear regression.
3. Rancangan aplikasi lanjut, Universitas Ichsan Gorontalo dengan judul “Prediksi Jumlah Mahasiswa Registrasi Per Semester Menggunakan Linier Regresi Pada Universitas Ichsan Gorontalo”, oleh Amiruddin dan Rezqiwati Ishak, penelitian ini mencoba melakukan prediksi jumlah mahasiswa yang registrasi di setiap semesternya dengan menggunakan metode Linear Regression yang bertujuan untuk mengetahui jumlah mahasiswa teregistrasi.
4. Rancangan aplikasi lanjut, Politeknik Kediri dengan judul “Prediksi Harga Saham Menggunakan Improved Multiple Linear Regression Untuk Pencegahan Data Outlier”, oleh Abidatul Izzah dan Ratna Widyastuti, penelitian ini dilakukan untuk memprediksi harga saham yang menggunakan metode multiple Linear Regression dan knn untuk meniadakan hasil prediksi mana yang paling mendekati.
5. Rancangan aplikasi lanjut, Universitas Mercu Buana Jakarta dengan judul “Analisis Performansi Algoritma Linear Regression Dengan Generalized Linear Model Untuk Prediksi Penjualan Pada Usaha Mikro, Kecil, Dan Menengah” oleh Fitria Habibatul Hamdanah dan Devi Fitriana, penelitian ini dilakukan untuk memprediksi penjualan pada usaha dalam lingkup mikro, kecil, dan menengah. Dengan tujuan untuk memprediksi penjualan barang umk di tahun berikutnya.

2. METODE PENELITIAN

2.1. Data Set

Dalam pembuatan aplikasi prediksi pengunjung wisatawan mancanegara bali, dengan menggunakan data set pengunjung yang di dapat kan dari situs web data pengunjung wisatawan bali 2019 dengan data set sebanyak 184 negara dengan jumlah pengunjung yang berbeda. Data yang di dapatkan berupa data excel, seperti pada gambar di bawah :

Gambar 1. data set mentah.

	Negara	JAN019	FEB019	MAR019	APR019	MEO19	JUN019	JUL019	AGT019	SEP019	—	MAR200	APR200	MAY200	JUNE200	JULY200	AUG200	S
0	Australian	95536	67474	83092	90265	98625	110815	123293	118556	124415	—	40039	7	0	0	0	0	1
1	Bhutan	60	30	47	38	31	31	46	67	43	—	20	0	0	0	0	0	0
2	Bangladesh	2262	2094	2489	2264	1938	4260	2656	2307	2035	—	568	2	0	0	0	0	0
3	Fiji	91	73	72	77	42	62	2656	80	68	—	32	0	0	0	0	0	0
4	Hongkong	1313	1367	764	6	1	5	2	1	0	—	1	0	0	0	0	0	0

5 rows x 25 columns

Gambar 2. Data set setelah di olah.

Dari data set yang sudah didapatkan akan di proses dengan menggunakan metode linear regression yang bertujuan untuk mendapatkan nilai R^2 yang dapat menghasilkan angka ketepatan terhadap prediksi pengunjung wisatawan bali dari dua bulan perbandingan.

2.2. Algoritma Pembangunan Sistem

Pembangunan sistem prediksi pengunjung wisatawan bali dengan data set berupa data pengunjung mancanegara yang berkunjung wisata ke bali menggunakan metode *Linear Regression*. regresi linear merupakan pendekatan untuk memodelkan hubungan antara suatu (satu atau lebih) variabel dependen dengan satu (regresi linear sederhana) atau lebih variabel independen (regresi linier banyak).

Salah satu aplikasi dari regresi linier adalah untuk melakukan prediksi berdasarkan data-data yang telah dimiliki sebelumnya. Dengan asumsi hubungan di antara variabel-variabel tersebut dapat didekati oleh suatu persamaan garis lurus, maka model yang mendekati hubungan antar variabel di data tersebut disebut sebagai model regresi linear.

Linear regression adalah alat statistik yang dipergunakan untuk mengetahui pengaruh antara satu atau beberapa variabel terhadap satu buah variabel. Variabel yang mempengaruhi sering disebut variabel bebas, variabel independen atau variabel penjelas. Dengan *Linear Regression* dalam melakukan prediksi hanya memiliki satu perubahan untuk populasi adalah (1).

$$Y = a + bx \quad (1)$$

Y = Subyek dalam variabel dependen yang diprediksikan.

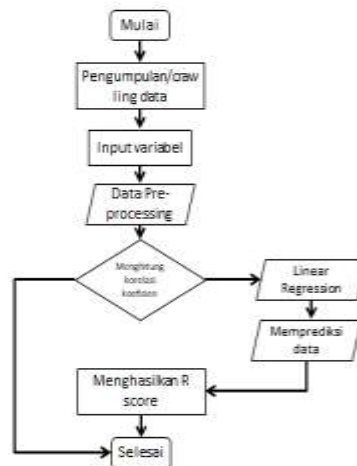
X = Subyek pada variabel independen yang mempunyai nilai tertentu.

a = Parameter intercept.

b = Parameter koefisien regresi variabel bebas.

2.3. Perancangan Aplikasi

Sebelum membuat aplikasi tentunya merancang terlebih dahulu sistem yang akan dibangun, hal ini dimaksudkan agar sistem berjalan sesuai alur yang sudah di bentuk. Penjelasan alur rancangan sistem dapat dilihat pada flow chart di bawah ini.



Gambar 3. Flochart Program

2.3. Pengumpulan Data

Pengumpulan data dilakukan dengan mengambil data pengunjung wisatawan bali pada situs web yang menyediakan data pengunjung wisatawan bali, dari total data pengunjung wisatawan bali ada sebesar 148 negara yang terdaftar pernah berwisata ke bali. Dari data tersebut akan dipilih input untuk sebagai variabel yang akan di prediksi menggunakan metode *Linear Regression*.

Variabel tersebut merupakan variabel negara sebagai input prediksi pengunjung wisatawan bali. Variabel ini kemudian yang akan menentukan nilai prediksi terhadap jumlah pengunjung wisatawan bali dari prediksi yang akan memakai dua bulan sebagai parameter perhitungan nilai r^2 pada proses regresi menggunakan *Linear Regression*.

2.4. Menghitung Korelasi dan Koefisien

Setelah dilakukan data pre-processing kemudian data masuk ke proses perhitungan korelasi dan koefisien untuk menentukan nilai r_1 , r_2 , dan r_3 jika mendapatkan hasil yang sesuai yaitu nilai nya linear maka proses akan berlanjut jika tidak maka kemungkinan ada yang salah dalam data atau data belum siap di jadikan data set.

Koefisien dan korelasi merupakan hubungan antara variabel satu dan lain nya. Koefisien korelasi dilambangkan dengan r , apabila koefisien korelasi mendekati -1 atau +1 maka dapat diartikan hubungan antar variabel tersebut kuat. Jika koefisien determinasi $R = 0$ maka dapat diartikan variabel bebas yang tidak ada kaitan nya dengan variabel terkait.

2.5. Model Linear Regression

Pendekatan untuk memodelkan hubungan antara suatu (satu atau lebih) variabel dependen dengan satu (regresi linear sederhana) atau lebih variabel independen (regresi linier banyak). Salah satu aplikasi dari regresi linier adalah untuk melakukan prediksi berdasarkan data-data yang telah dimiliki sebelumnya. Dengan asumsi hubungan di antara variabel-variabel tersebut dapat didekati oleh suatu persamaan garis lurus, maka model yang mendekati hubungan antar variabel di data tersebut disebut sebagai model regresi linear.

Pada tahap ini setelah menghitung *linear regression* nya maka akan menghasilkan nilai *regression coefficient beta* dan *intercept coefficient alpha* untuk menentukan nilai terbaik pada R^2 di tahap selanjut nya.

2.6. Metode Evaluasi

Metode evaluasi yang akan digunakan dalam pembuatan program ini adalah dengan menghitung nilai R^2 koefisien determinasi. Nilai R^2 menunjukkan persentase variabel tak bebas dapat dijelaskan oleh variabel bebas. Nilai koefisien determinasi adalah $0 < R^2 < 1$. Semakin tinggi nilai R^2 maka semakin baik model karena semakin besar keragaman peubah dependen yang dapat dijelaskan oleh peubah independen. Perhitungan R^2 dapat dilakukan dengan persamaan (2)

$$R^2 = 1 - \frac{SS_{error}}{SS_{total}} = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \quad (1)$$

Keterangan :

y_i = Observasi respons ke-i

$\hat{y}_i$ = Ramalan respons ke-i

$\bar{y}$ = rata-rata

3. HASIL DAN PEMBAHASAN

3.1. Hasil Penelitian

3.1.1 Load jumlah pengunjung Bali dataset using pandas

Mengetahui kolom - kolom yang ada di dalam dataset, Terlebih dahulu menentukan kelas target yang ingin diprediksi yaitu jumlah pengunjung Bali. Setelah mengetahui kelas target yang ingin prediksi, kemudian menentukan fitur-fitur apa saja yang akan kami gunakan dalam melakukan prediksi.

Tabel 1

```
bali.head()
```

	Negara	JAN019	FEB019	MAR019	APR019	MAY019	JUN019	JUL019	AGT019	SEP019
0	Australian	99556	87474	83002	90295	98625	110815	123293	118558	124416
1	Shutan	60	30	47	38	31	31	46	67	43
2	Bangladesh	2262	2094	2409	2204	1938	4265	2856	2307	2005
3	Fij	91	73	72	77	42	62	2956	80	68
4	Hongkong	1313	1367	764	6	1	5	2	1	0

5 rows x 25 columns

3.1.2 Penjelasan Data

Berikut adalah fitur-fitur yang kami pilih beserta tipe datanya sesuai dengan table 2.

Tabel 1. Info Data

```
bali.info()
```

#	Column	Non-Null Count	Dtype
0	Negara	184 non-null	object
1	JAN019	184 non-null	int64
2	FEB019	184 non-null	int64
3	MAR019	184 non-null	int64
4	APR019	184 non-null	int64
5	MAY019	184 non-null	int64
6	JUN019	184 non-null	int64
7	JUL019	184 non-null	int64
8	AGT019	184 non-null	int64
9	SEP019	184 non-null	int64
10	OKT019	184 non-null	int64
11	NOV019	184 non-null	int64
12	DES019	184 non-null	int64
13	JAN200	184 non-null	int64
14	FEB200	184 non-null	int64
15	MAR200	184 non-null	int64
16	APR200	184 non-null	int64
17	MAY200	184 non-null	int64
18	JUNE200	184 non-null	int64
19	JULY200	184 non-null	int64
20	AUG200	184 non-null	int64
21	SEP200	184 non-null	int64
22	OCT200	184 non-null	int64
23	NOV200	184 non-null	int64
24	DEC200	184 non-null	int64

dtypes: int64(24), object(1)
memory usage: 36.1+ KB

3.1.3 Describe dataset

Pada Tabel 2 dapat dilihat menggunakan describe sudah menampilkan hasil count, mean, std, min, 25%, 50%, 75%, max. Dalam Scatter Plot ini, dapat dilihat jarak plot titik nilai harga prediksi tidak sangat jauh dengan jumlah pengunjung, ini menyatakan mempunyai hasil yang akurat Hasil model terbaik tersebut akan nanti dibandingkan dengan algoritma.

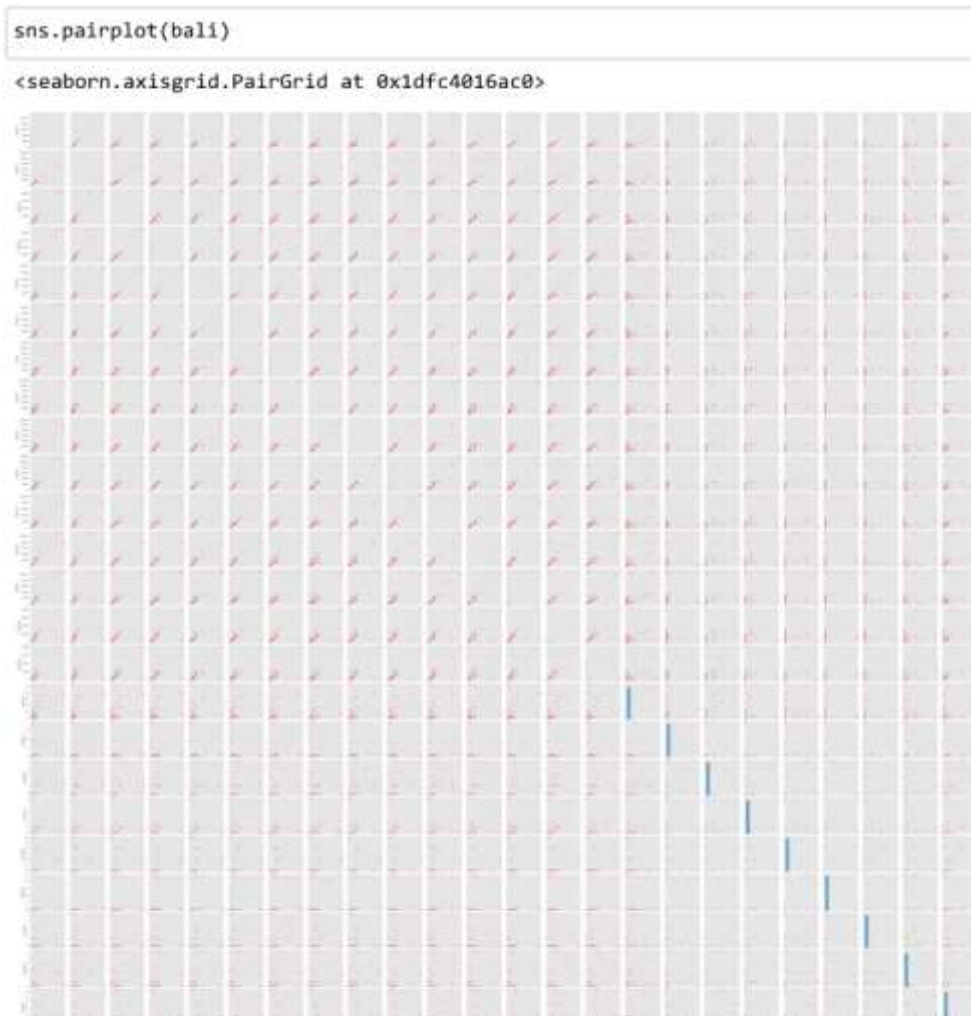
Tabel 3. Penjelasan isi dalam data

bali.describe()						
	JAN019	FEB019	MAR019	APR019	MAY019	JUN019
count	184.000000	184.000000	184.000000	184.000000	184.000000	184.000000
mean	2463.538043	2376.853261	2436.233696	2576.895652	2638.565217	2967.945652
std	11397.676884	10841.315106	9729.992146	10328.860638	10914.362896	12138.378496
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	7.750000	8.000000	8.500000	9.750000	8.000000	9.000000
50%	76.500000	67.000000	65.000000	77.000000	49.000000	87.000000
75%	759.250000	666.750000	793.250000	749.750000	646.000000	596.000000
max	113776.000000	122643.000000	89178.000000	94202.000000	98625.000000	110815.000000

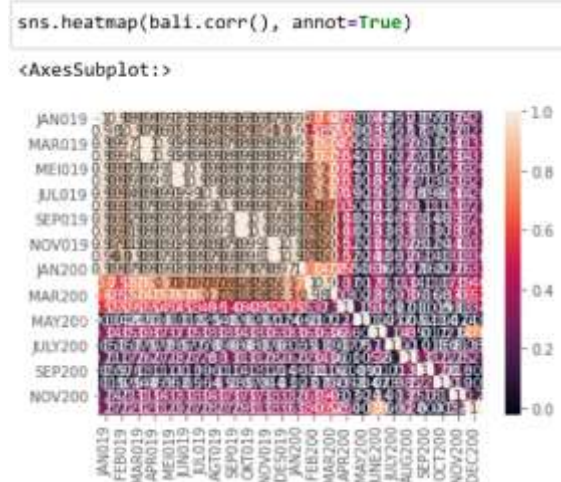
8 rows x 24 columns

3.1.4 Exploratory Data Analysis

Tabel 4. Diagram plot dari data perbulan.



Tabel 5



3.1.5 Train simple linear regression

Fungsi utamanya adalah untuk memisahkan data sebagai “train” dan “test.”. Data keseluruhan akan dibagi menjadi 80% sebagai *train* dan 20% sebagai tes. “value-y” akan menjadi “median_JAN_value”, dan “value_x” akan menjadi “penghasilan_median”.

```
X = bali['JAN019']
y = bali['JAN200']
```

```
from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3,
```

Gambar 4. Menghitung x dan y.

```
from sklearn import metrics
from sklearn.model_selection import cross_val_score

def cross_val(model):
    pred = cross_val_score(model, X, y, cv=10)
    return pred.mean()

def print_evaluate(true, predicted):
    mae = metrics.mean_absolute_error(true, predicted)
    mse = metrics.mean_squared_error(true, predicted)
    rmse = np.sqrt(metrics.mean_squared_error(true, predicted))
    r2_square = metrics.r2_score(true, predicted)
    print('MAE:', mae)
    print('MSE:', mse)
    print('RMSE:', rmse)
    print('R2 Square', r2_square)
    print('_____')

def evaluate(true, predicted):
    mae = metrics.mean_absolute_error(true, predicted)
    mse = metrics.mean_squared_error(true, predicted)
    rmse = np.sqrt(metrics.mean_squared_error(true, predicted))
    r2_square = metrics.r2_score(true, predicted)
    return mae, mse, rmse, r2_square
```

Gambar 5. Menghitung nilai R square.

3.1.6 Calculate the pearson correlation coefficient

```
x = (bali["JAN019"]-np.mean(bali["JAN019"]))
y = (bali["JAN200"]-np.mean(bali["JAN200"]))
r1 = np.sum(x*y)/(bali.shape[0]-1)/np.std(bali["JAN019"], ddof=1)/np.std(bali["JAN200"], ddof=1)
r2 = stats.pearsonr(bali["JAN019"], bali["JAN200"])
r3 = np.corrcoef(bali["JAN019"], bali["JAN200"])
print("R1 = %.3f" % (r1))
print("R2 = %.3f" % (r2))
print("R3 = %.3f" % (r3[0][1]))
```

R1 = 0.995
R2 = 0.995
R3 = 0.995

Gambar 6. Hasil coefficient.

Ada hubungan yang kuat antara skor pengetahuan dengan asupan kalsium Model regresi harus dekat dengan liner.

3.1.7 Report the regression and intercept coefficients

```
regr = LinearRegression()
regr.fit(np.array(X_train).reshape(-1,1), y_train)
```

LinearRegression()

regr.intercept_
213.3419218938875

regr.coef_
array([1.10765946])

regr.intercept_, regr.coef_
(213.3419218938875, array([1.10765946]))

Gambar 7. Hasil Regression.

3.1.8 evaluation

```
regrq = regr.predict(X)
print('R2 (sklearn):', r2_score(bali["JAN019"], regrq))

ssT = [ ((bali["JAN019"][i]-np.mean(bali["JAN019"]))**2) for i in range(0, len(bali["JAN019"]))]
ssR = [ ((bali["JAN019"][i]-regrq[i])**2) for i in range(0, len(bali["JAN019"]))]
ssM = [ ((regrq[i]-np.mean(bali["JAN019"]))**2) for i in range(0, len(bali["JAN019"]))]

print("R2 (met A)=", np.sum(ssM)/np.sum(ssT))
print("R2 (met B)=", 1 - np.sum(ssR)/np.sum(ssT))
```

n=40
k=2
adj_r2_score = 1 - ((1-r2)*(n-1)/(n-k-1))
print(adj_r2_score)

0.9944073090361236

Gambar 8.

	Model	MAE	MSE	RMSE	R2 Square	Cross Validation
0	Linear Regression	81135.566093	1.006842e+10	100341.529545	0.914682	0.917379
1	Robust Regression	84645.310693	1.099681e+10	104865.656302	0.906815	0.911225
2	Ridge Regression	81428.648355	1.015327e+10	100763.435337	0.913963	0.917379

Gambar 9. Hasil MAE, MSE, RMSE

3.1.9 Predictions from our Model

Hasil prediksi jumlah pengunjung bali.

```
pred = regr.predict(np.array(X_test).reshape(-1,1))
y_test.head()
```

```
19      11
42       0
156     502
111       8
148    1618
Name: JAN200, dtype: int64
```

Gambar 9.

```
pd.DataFrame({'True Values': y_test, 'Predicted Values': pred})
```

	True Values	Predicted Values
19	11	224.418516
42	0	215.557241
156	502	570.008268
111	8	231.004473
148	1618	1496.226090
15	752	862.889533
24	6	219.987879
68	7	237.710430
117	884	836.954198
98	48	230.925749
97	16	218.880219
162	3834	1162.606003
69	63	265.401917
174	1493	1075.100002
45	25	226.633633
18	112954	126236.404700
51	1	213.341922

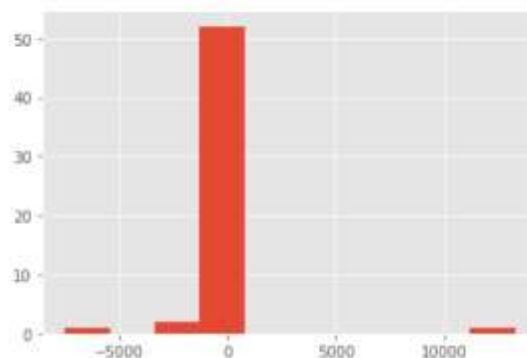
Gambar 10.

3.1.10 Residual Histogram

Ini akan menunjukkan seberapa jauh nilainya. Menunjukkan nilai prediksi dikurangi nilai pengujian sebenarnya untuk semua titik data. Kemudian, kita harus memplot dengan histogram untuk melihat seberapa "off" setiap nilai. Ini dapat dilakukan dengan perintah berikut

```
residuals = pred - y_test
plt.hist(residuals)

(array([ 1.,  0.,  2., 52.,  0.,  0.,  0.,  0.,  0.,  1.]),
 array([-7542.04292581, -5459.3981632 , -3376.75340058, -1294.10863796,
        788.53612466, 2871.18088728, 4953.8256499 , 7036.47041252,
        9119.11517514, 11201.75993776, 13284.40470038]),
 <BarContainer object of 10 artists>)
```



Gambar 11. Grafik di atas menunjukkan distribusi kesalahan.

4. KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan hasil pencarian data pengunjung wisatawan bali dengan menggunakan web, di dapat data set pengunjung wisatawan bali pada tahun 2019 dengan data set sebanyak 184 negara dengan jumlah pengunjung yang berbeda. Pembangunan sistem prediksi pengunjung wisatawan bali dengan menggunakan metode Linear Regression sangat membantu dan memudahkan dalam memprediksi jumlah pengunjung wisatawan bali. dapat diambil kesimpulan :

1. Setelah melakukan pengujian, variabel yang mempunyai pengaruh besar terhadap data pengunjung pada setiap bulan nya. terjadi pada tabel 3.1
2. Dari hasil pengujian dataset menggunakan evaluation matriks dengan parameter yang berbeda, kami menemukan bahwa model evaluation matriks dengan jumlah MAE 81135.566093, MSE = 1.0006842, RMSE = 100341.529545. Menghasilkan hasil prediksi dengan tingkat akurasi R^2 0.994407.

Oleh karena itu program ini sangat membantu untuk menentukan jumlah pengunjung di tahun beikut nya untuk yang memiliki usaha dan pengelola dapat mempersiapkan semua yang diperlukan. melihat dari hasil program yang sudah dibuat memiliki hasil pengujian.

Berdasarkan hasil penelitian dapat diperkirakan sekitar 941.365 pengunjung dari kelima Negara yang berbeda pada tahun 2019. Pembangunan sistem prediksi pengunjung wisatawan bali dengan menggunakan metode Linear Regression sangat membantu dan memudahkan dalam memprediksi jumlah pengunjung wisatawan bali.

Saran

Saran untuk penelitian selanjutnya agar dapat menggunakan metode selain dari Linear Regresi.

DAFTAR PUSTAKA

- [1] Darwin Prasetyo and Dra. Harlili, M.Sc., “*Predicting Football Match Results with Logistic Regression*,” 2016.
- [2] Roseline O. Ogundokun, Adewale F. Lukman, Golam B.M. Kibria, Joseph B. Awotunde and Benedita B. Aladeitan, “*Predictive modelling of COVID-19 confirmed cases in Nigeria*”, R.O. Ogundokun et al. / Infectious Disease Modelling 5 (2020).
- [3] Amiruddin dan Rezqiwati Ishak, “Prediksi Jumlah Mahasiswa Registrasi Per Semester Menggunakan Linier Regresi Pada Universitas Ichsan Gorontalo”, Universitas Ichsan Gorontalo. ILKOM Jurnal Ilmiah Volume 10 Nomor 2 Agustus 2018.
- [4] Abidatul Izzah dan Ratna Widyastuti, “Prediksi Harga Saham Menggunakan *Improved Multiple Linear Regression* Untuk Pencegahan Data Outlier”, Politeknik Kediri, KINETIK, Vol. 2, No. 3, Agustus 2017.
- [5] Fitria Habibatul Hamdanah dan Devi Fitriana, “Analisis Performansi Algoritma *Linear Regression* Dengan Generalized Linear Model Untuk Prediksi Penjualan Pada Usaha Mikro, Kecil, Dan Menengah”, Maret 2021.