



## PENERAPAN DATA MINING UNTUK MEMPREDIKSI PENJUALAN KAIN TENUN MNGGUNAKAN REGRESI LINEAR

### STUDI KASUS: UD.BINTANG REMAWE SUKARARE

Firda Widiastuti<sup>a</sup>, Wafiah Murniati<sup>b</sup>, Saikin<sup>c</sup>

<sup>a</sup>Teknik Informatika, [firdawidia46@gmail.com](mailto:firdawidia46@gmail.com), STMIK Lombok

<sup>b</sup>Teknik Informatika, [wafiah.mr@gmail.com](mailto:wafiah.mr@gmail.com), STMIK Lombok

<sup>c</sup>Teknik Informatika, [eken.apache@gmail.com](mailto:eken.apache@gmail.com), STMIK Lombok

#### Abstract

Weaving craft is one of the handicrafts located in Lombok, West Nusa Tenggara. Weaving is one of the MSMEs that is very close to the tourism industry and has good economic potential because it absorbs a lot of labor, opens up business fields, and increases the country's foreign exchange. The problem that is often faced by entrepreneurs of woven fabrics is the difficulty in estimating customer demand, so that some of the products requested by customers are not available. It is necessary for sales analysis of woven fabric products to be able to predict customer demand, by means of analyzing past sales data to predict future sales. The research that will be carried out is to predict sales of woven fabric products by processing sales data in the past by modeling the Linear Regression method, and for testing the algorithm is by Mean Square Error (MSE), Mean Square Error (MSE) and Root Mean Square Error. (RMSE). From the results of linear regression modeling the score obtained is 0.8041320270845731. and the test results mean Mean Square Error (MSE) the error value obtained is too high, namely 47,377, and the Root Mean Square Error (RMSE) value is 6.883125, while the MAE score is 3.373572.

**Keywords:** MSME, Prediction, Linear Regression, Means Square Error (MSE)

#### Abstrak

Kerajinan Tenun adalah salah satu kerajinan tangan yang berada di Lombok Nusa Tenggara Barat. Kerajinan Tenun merupakan salah satu UMKM yang sangat dekat dengan industri pariwisata dan memiliki potensi ekonomi yang baik karena menyerap banyak tenaga kerja, membuka lapangan usaha, dan menambah devisa negara. Permasalahan yang sering dihadapi oleh pengusaha kain tenun sulitnya mengestimasi permintaan pelanggan, sehingga beberapa produk yang diminta oleh pelanggan tidak tersedia. Diperlukan untuk analisis penjualan produk kain tenun untuk dapat memprediksi permintaan pelanggan, dengan cara analisis data penjualan masa lalu untuk memprediksi penjualan berikutnya. Pada penelitian yang akan dilakukan ialah akan melakukan prediksi penjualan produk kain tenun dengan mengolah data penjualan dimasa lalu dengan melakukan pemodelan metode Regresi Linier, dan untuk pengujian algoritma ialah dengan Mean Square Error (MSE), Mean Square Error (MSE) dan Root Mean Square Error (RMSE). Dari hasil pemodelan regersi linear score yang didapatkan sebesar 0.8041320270845731. dan hasil pengujian mean Mean Square Error (MSE) nilai error yang didapatkan.terlalu tinggi yakni sebesar 47.377, dan nilai Root Mean Square Error (RMSE) sebesar 6.883125, Sedang score MAE sebesar 3.373572.

**Kata kunci:** UMKM, Prediksi, Regresi Linear, Means Square Error (MSE)

## 1. PENDAHULUAN

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik-teknik, metode-metode, atau algoritma dalam data mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses Knowledge Discovery in Database (KDD) secara keseluruhan. (Mardi, 2017).

Salah satu fungsi data mining yang sering digunakan adalah untuk memprediksi. Nilai dari hasil prediksi akan digunakan dimasa yang akan datang berdasarkan data-data sebelumnya (Efori Buulolo(2020)). Peramalan atau prediksi memiliki fungsi untuk mengestimasi nilai dari sebuah atau kumpulan variabel dimasa yang akan datang menggunakan data masa lalu dari nilai tersebut (Efendi (2018)). Berdasarkan fungsinya, peramalan dibagi menjadi 3, salah satunya yaitu Peramalan Permintaan (Ardiansa (2018)).

Kerajinan Tenun adalah salah satu kerajinan tangan yang berada di Lombok Nusa Tenggara Barat. Kerajinan Tenun merupakan salah satu UMKM yang sangat dekat dengan industri pariwisata dan memiliki potensi ekonomi yang baik karena menyerap banyak tenaga kerja, membuka lapangan usaha, dan menambah devisa Negara. Produk kerajinan tangan kain tenun yang dihasilkan selain populer dipasar domestik juga menembus pasar manca negara.

Permasalahan yang sering dihadapi oleh pengusaha kain tenun sulitnya mengestimasi permintaan pelanggan, sehingga beberapa produk yang diminta oleh pelanggan tidak tersedia. Begitu juga dialami oleh UD. BINTANG REMAWE, merupakan Usaha dagang yang menjual produk kain tenun yang beralamatkan di Jln. Tenun Sukarara, Lombok Tengah, NTB. Setelah melakukan observasi ada beberapa permasalahan yang dihadapi oleh usaha dagang tersebut, seperti sulit memprediksi permintaan pelanggan, permintaan pelanggan yang tidak bisa dipenuhi karena ketersediaan barang yang kurang dan barang-barang yang kurang laku banyak menumpuk digudang.

Berdasarkan permasalahan tersebut, diperlukan untuk analisis penjualan produk kain tenun untuk dapat memprediksi permintaan pelanggan, dengan cara analisis data penjualan masa lalu untuk memprediksi penjualan berikutnya. Data mining ialah salah satu bidang ilmu yang menerapkan pola analisis data masa lalu untuk memprediksi yang akan terjadi dimasa depan dengan menerapkan pembelajaran terhadap data. Secara garis besar data mining terdiri dari beberapa metode analisis, yakni analisis klasifikasi, analisis regresi dan analisis cluster. dan salah satu yang sering digunakan ialah untuk prediksi ialah analisis regresi. Regresi Linear sudah banyak diterapkan dalam prediksi, salah satunya adalah prediksi penjualan buku (Kamal,2017). Regresi Linear merupakan salah satu metode yang terdapat dalam data mining. Regresi linear juga diterapkan untuk prediksi penjualan properti. Prediksi penjualan properti menggunakan metode regresi linear ini dapat dikatakan tergolong dalam kategori sangat baik didasari hasil pengujian keakurasian terhadap keseluruhan tipe yang menampilkan nilai MSE, RMSE, dan MAPE yang memenuhi standar (Ayuni,2019)

Pada penelitian yang akan dilakukan ialah akan melakukan prediksi penjualan produk kain tenun dengan mengolah data penjualan dimasa lalu dengan menggunakan metode Regresi Linier. Selain melakukan prediksi prosesnya ini akan ditekankan pada proses preprocessing dan pengujian hasil regresi linier dengan Mean Square Error (MSE). Data yang digunakan merupakan data proses penjualan produk kain tenun dan merupakan data beberapa bulan terakhir.

## 2. TINJAUAN PUSTAKA

### 2.1. Tinjauan Pustaka

**2.1.1.** Muhartini (2021), dalam penelitiannya yang berjudul “Analisis Peramalan Jumlah Penerimaan Mahasiswa Baru Dengan Menggunakan Metode Regresi Linear Sederhana” bertujuan melakukan perkiraan untuk mengetahui banyaknya jumlah mahasiswa baru pada Prodi Manajemen Fakultas Ekonomi dan Bisnis menggunakan metode regresi sederhana linear. Sehingga dapat diketahui jumlah mahasiswa baru dimasa yang akan datang. Penelitian ini menggunakan data selama 5 tahun terakhir yaitu dari tahun

2015 sampai tahun 2019. Dengan Hasil prediksi jumlah penerimaan mahasiswa baru dari prodi Manajemen dengan hasil pengujian tingkat error sebesar 3,444% atau tingkat akurasi 96,556%. [4]

**2.1.2.** Trianggana (2020) dalam penelitiannya “Peramalan Jumlah Siswa-Siswi Melalui Pendekatan Metode Regresi Linier” Aplikasi ini dapat meramalkan data untuk tahun pelajaran berikutnya berdasarkan analisa data pada tahun pelajaran sebelumnya terhadap jumlah pendaftara siswa baru per jurusan. Berdasarkan pengujian yang telah dilakukan, dapat disimpulkan bahwa fungsional dari aplikasi telah berjalan sebagaimana mestinya dan aplikasi mampu menampilkan hasil peramalan berdasarkan data trend yang diinputkan. [5]

**2.1.3.** Hendra (2020), Penelitiannya yang berjudul ” Penerapan Data Mining Untuk Prediksi Penjualan Readymix Menggunakan Metode Algoritma C4.5 Pada Pt Remicon Widyaprim” Dalam melakukan analisis peneliti menggunakan teknik data mining dengan menggunakan algoritma klasifikasi c4.5 yang dilakukan dengan menggunakan perhitungan manual dan juga peneliti menggunakan aplikasi weka versi 3.9.4 untuk mengetahui hasil dari prediksi penjualan readymix di PT remicon widyaprima. dari hasil penelitian ini, data yang diperoleh dapat bekerja secara akurat pada aturan setiap tes, ini dapat memudahkan penelitian berbasis pada data yang telah diperoleh melalui proses pengujian prediksi penjualan yang dikategorikan dalam beberapa variabel. [6]

Perbedaan Penelitian ini dengan Penelitian sebelumnya adalah dalam penelitian yang akan dilakukan, yang berlatar belakang dengan permasalahan yang ada pada UMKM di UD. Bintang Remawe, Sukarara, Lombok Tengah atau tempat penelitian berlangsung, peneliti bertujuan untuk memprediksi penjualan berikutnya dengan menggunakan metode regresi linear. Karena berdasarkan beberapa tinjauan pustaka yang telah dicantumkan dan menjadi acuan bagi peneliti, metode regresi linear terbukti dapat memberikan solusi dalam peramalan. Data yang diambil dalam penelitian ini menggunakan data penjualan pada juni 2020 sampai dengan juni 2021. Dengan harapan mendapatkan hasil prediksi agar para pelaku UMKM dapat mengelola usaha dengan baik berdasarkan hasil prediksi pada penjualan bulan berikutnya.

## **2.2. Dasar Teori**

### **2.2.1. Usaha Mikro Kecil Menengah (UMKM)**

UMKM adalah singkatan dari Usaha Mikro Kecil dan Menengah, UMKM artinya sebagai bisnis yang dijalankan individu, rumah tangga, atau badan usaha ukuran kecil lainnya[7].

### **2.2.2. Kerajinan Tangan Kain Tenun**

Kain Tenun Songket Lombok adalah kain tenun yang memiliki hiasan timbul dari benang katun, benang perak atau benang emas.[8].

### **2.2.3. Data Mining**

Data Mining merupakan suatu proses ekstraksi atau penggalian data informasi yang besar, yang belum diketahui sebelumnya, namun dipahami dan berguna dari database yang besar untuk membuat suatu keputusan bisnis yang sangat penting [9].

### **2.2.4. Reprocessing Data**

*Reprocessing* data merupakan teknik awal untuk mengubah data mentah atau biasa dikenal dengan *raw* data yang dikumpulkan dari berbagai sumber menjadi informasi yang lebih bersih dan bisa digunakan untuk pengolahan selanjutnya[10].

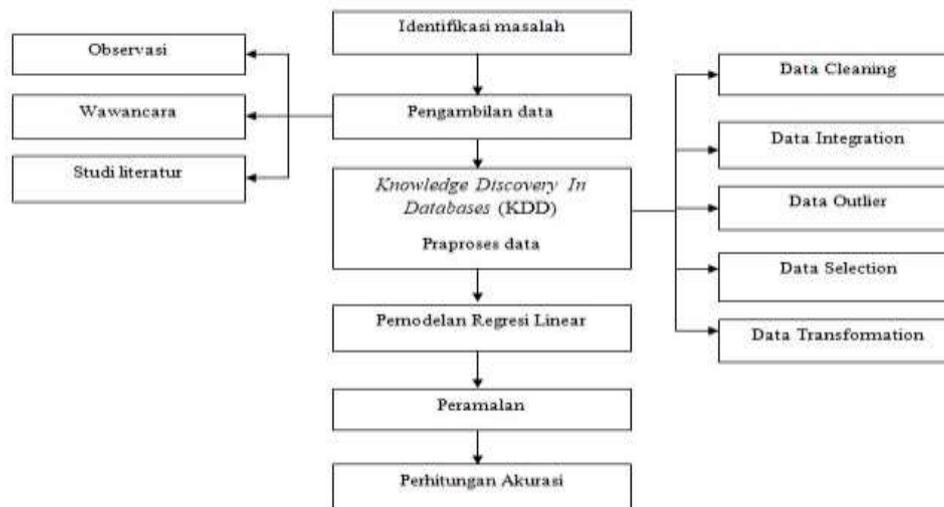
### **2.2.5. Regresi Linear**

*Regresi* Linear adalah Metode statistik yang berfungsi untuk menguji sejauh mana hubungan sebab akibat antara variabel faktor penyebab (X) terhadap Variabel akibatnya[11].

## **3. METODOLOGI PENELITIAN**

---

*PENERAPAN DATA MINING UNTUK MEMPREDIKSI PENJUALAN KAIN TENUN  
MNGGUNAKAN REGRESI LINEAR STUDI KASUS: UD.BINTANG REMAWE SUKARARE  
(Firda Widiastutia)*



Gambar 1. Diagram Alir Penelitian

### 3.1. Identifikasi Masalah

Kain tenun songket songket merupakan salah satu produk UMKM yang produksi terbesarnya ada dibabupaten lombok tengah. Berdasarkan studi lapangan pendahuluan yang dilakukan di UD. BINTANG REMAWA Desa Sukarara, Lombok Tengah, NTB., didapatkan berbagai permasalahan diantaranya ialah sulitnya memprediksi penjualan produk setiap harinya. Selain itu, permintaan konsumen yang tidak menentu seperti permintaan yang mengalami peningkatan. Dengan adanya peningkatan transaksi penjualan tersebut, para pelaku UMKM harus merencanakan dan menyiapkan persediaan bahan baku berikutnya karna dalam proses pembuatan kain tenun ini memakan waktu yang cukup lama sehingga membuat para pelaku UMKM kesulitan dalam memenuhi permintaan barang.

Oleh karena itu, Analisis Data penjualan kain tenun diperlukan untuk menyelesaikan masalah tersebut. Analisa data yang digunakan adalah data mining dengan metode regresi linear.

### 3.2. Pengambilan Data

3.2.1. Observasi

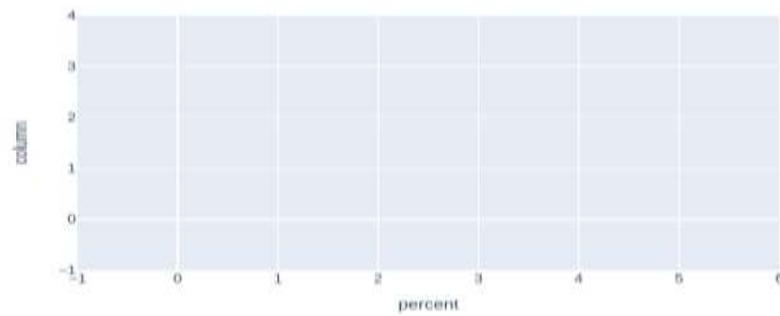
3.2.2. Wawancara

3.2.3. Studi Literatur

### 3.3. Knowledge Discovery in Database (KDD)

#### 3.3.1. Data Cleaning

Membersihkan data yang missing ada jenis langkah yang digunakan ialah dengan menghapus fitur yang nilai missingnya yang terlalu tinggi dan menginputkan dengan nilai rata-rata atau mean pada fitur yang nilai missingnya sedikit.



Gambar 3.1 Pembersihan Data

### 3.3.2 Data Integration

Data yang didapat berupa file-file yang terpisah dalam bentuk data penjualan perbulan maka perlu dilakukan integrasi data pada file data penjualan tersebut. tampilan sourcode dan tampilan bentuk setelah diintegrasikan seperti gambar dibawah.

```
In [3]: path="./data penjualan"
files=[file for file in os.listdir(path) if not file.startswith('.')]
df=pd.DataFrame()

for file in files:
    current_data=pd.read_excel(path+"/"+file)
    df=pd.concat([df, current_data])
df.to_csv("all_data_copy.csv", index=False)
print(path, "\nFILES =", files)
print

./data penjualan :
FILES = ['september.xls', 'mei.xls', 'agustus.bak', 'juni.bak', 'oktober.xls', 'april.xls', 'agustus.xls', 'juni.xls',
'januari.xls', 'pebruari.xls', 'juli.xls', 'maret.xls']

Out[3]: <function print>

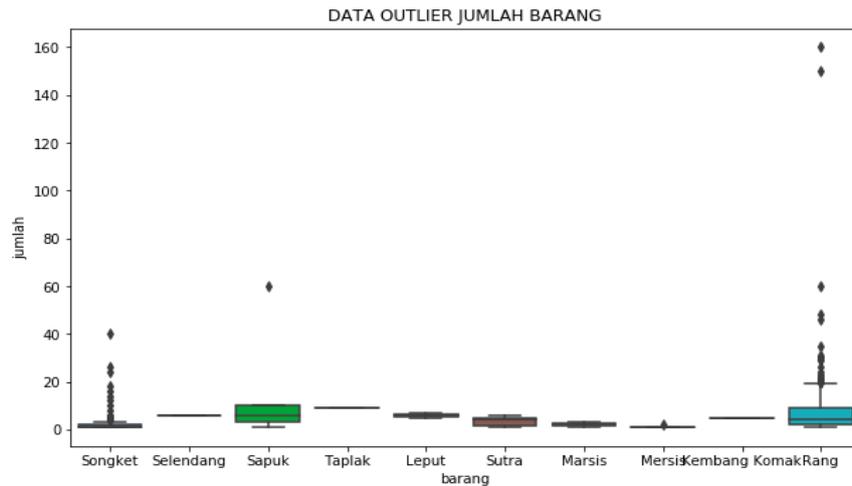
In [4]: data=pd.read_csv("all_data_copy.csv")
data.shape

Out[4]: (508, 8)
```

Gambar 3.2 Data Integreration

### 3.3.3 Data Otlir (Pencilan Data)

Pencilan data adalah dimana posisi nilai dari data mempunyai rentang yang sangat jauh dengan kelompok data lainnya.



Gambar 3.3 Data Pencilan

**3.3.4 Normalisasi Data**

Nilai dari atribut yang berbeda-beda seperti kasus pencilan pada setiap fitur yang telah dianalisis terdapat sebagian kecil nilai yang rentangnya berjauhan dengan kelompok nilai lainnya.

	date	barang	jumlah	hbeli	hjual	totalharga
0	2020-01-01	Songket	2	0.087860	0.152941	0.152941
1	2020-01-01	Songket	1	0.011849	0.015686	0.015686
2	2020-02-01	Songket	4	0.012967	0.054902	0.054902
3	2020-02-01	Songket	1	0.016320	0.019608	0.019608
4	2020-02-01	Songket	3	0.012296	0.054902	0.054902

Gambar 3.4 Normalisasi Data

**3.4. Implementasi Prediksi Metode Regresi Linear**



Gambar: 3.5 Penyelesaian dengan Regresi Linear

#### 4. HASIL DAN PEMBAHASAN

##### 4.1. Pemodelan Regresi

##### 4.1.1. Dataset

Tabel 4.1 Dataset Penjualan Produk.

date	barang	jumlah	hbeli	hjual	totalharga
2020-01-01 00:00:00	7	2	0,08786	0,15294	0,15294
2020-01-01 00:00:00	7	1	0,01185	0,01569	0,01569
2020-02-01 00:00:00	7	4	0,01297	0,0549	0,0549
2020-02-01 00:00:00	7	1	0,01632	0,01961	0,01961
2020-02-01 00:00:00	7	3	0,0123	0,0549	0,0549
2020-03-01 00:00:00	6	6	0,00179	0,01961	0,01961
2020-04-01 00:00:00	7	1	0,05209	0,07451	0,07451
2020-04-01 00:00:00	7	1	0,00514	0,00588	0,00588
2020-05-01 00:00:00	7	1	0,05209	0,07451	0,07451
2020-07-01 00:00:00	7	3	0,00514	0,02353	0,02353
2020-09-01 00:00:00	7	4	0,05209	0,32941	0,32941
2020-12-01 00:00:00	7	1	0,01185	0,0098	0,0098
2020-01-14 00:00:00	7	1	0,01185	0,0098	0,0098
2020-01-14 00:00:00	7	1	0,03197	0,04314	0,04314
2020-01-15 00:00:00	5	10	0,00067	0,02549	0,02549
2020-01-16 00:00:00	7	1	0,00961	0,0098	0,0098
dst....	...	...	...	...	...

Pada tabel dataset diatas pada fitur Date nilai dari fitur tersebut terdiri dari tahun, bulan, tanggal dan waktu, tipe data yang digunakan dalam fitur tersebut ialah tipe data objek, sedangkan pada implementasinya algoritma regresi hanya bisa mengolah data dalam bentuk tipe data integer maka dibutuhkan untuk mengubah tipe data tersebut kedalam tipe data integer.

##### 4.1.2. Menampilkan Korelasi Data



Gambar 4.1 Korelasi Fitur Pada Dataset

**4.2. Data Variabel X dan Y**

Regresi linear bertujuan untuk mengetahui pengaruh suatu variabel terhadap variabel lainnya.

```
In [46]: y= df['jumlah'].values #Dependent Features (Class)
x=df.drop(['jumlah'],axis=1) #Independent Features
```

Gambar 4.4 Proses Pemisahan Varibel X dan Variabel Y

**4.3. Menentukan Data Trining dan data Testing**

Pemisahan data latih (Data testing) dan data testing akan digunakan tiga skenario, yang bertujuan untuk mencari nilai score terbaik dari fitting model algoritma regresi yang diterapkan. scenario pemisahan data train dan test disajikan dalam tabel berikut:

Tabel 4.2 skenario pembagian data train dan testing

skenario	train	test
50 : 50	50 %	50%
70: 30	70 %	70%
80 :20	80 %	20%

**4.3.1. Data Trining dan testing 50:50**

```
In [31]: from sklearn.model_selection import train_test_split

x_train,x_test,y_train,y_test = train_test_split(x,y,test_size=0.5,random_state=45)
print ("X train: ", len(x_train))
print("X test: ", len(x_test))
print("y train: ", len(y_train))
print("y_test: ", len(y_test))

X_train: 207
X_test: 207
y_train: 207
y_test: 207
```

Gambar 4.2 Data Trining dan Testing 50:50

Dari proses pemisahan data train dan testing dengan pembagian 50 : 50 maka didapatkan hasil yaitu  $x_{train} : 20$  ,  $x_{test} : 207$  ,  $y_{train} : 207$  ,  $y_{test} : 207$ .

```
In [48]: from sklearn.linear_model import LinearRegression
model=LinearRegression()
model.fit(x_train, y_train)
model.score(x_test, y_test)

Out[48]: 0.7075627489946119
```

Gambar 4.3 Score Nilai Regresi Skenario Pertama.

Setelah melakukan fitting model dengan data train 50 % dan test 50% maka skor yang didapatkan sebesar 70 persen.

#### 4.3.2. Data Training dan Testing 70:30

```
In [49]: 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,random_state=45)
print("X train: ", len(x_train))
print("X test: ", len(x_test))
print("y_train: ", len(y_train))
print("y_test: ", len(y_test))

X_train: 289
X_test: 125
y_train: 289
y_test: 125
```

Gambar 4.4 Splitting data Training dan Testing 70:30

Dari proses splitting data dengan skenario dua yakni 70 persen untuk data training dan 30 persen untuk data testing. maka jumlah data yang didapat untuk data training 289 dan jumlah data untuk data testing.

```
In [50]: from sklearn.linear_model import LinearRegression
model2=LinearRegression()
model2.fit(x_train, y_train)
model2.score(x_test, y_test)

Out[50]: 0.8041320270845731
```

Gambar 4.5 Score Fitting Model Skenario Kedua

Dari Gambar 4.10 hasil pemodelan regresi dengan menggunakan data training 70 persen dan data testing 30 persen, maka skor yang didapatkan sebesar 0.8041320270845731. Ini berarti jumlah persentase skor yang didapatkan sebesar 80 persen.

#### 4.3.3. Data training dan Testing 80:20

```
In [51]: from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test = train_test_split(x,y,test_size=0.2,random_state=45)
print ("X train: ", len(x_train))
print("X test: ", len(x_test))
print("y_train: ", len(y_train))
print("y_test: ", len(y_test))
X_train: 331
X_test: 83
y_train: 331
y_test: 83
```

Gambar 4.6 Splitting data Train dan Testing 80:20

Dari Gambar 4.11 proses split data didapat dengan menggunakan skenario ketiga yakni data train 80 persen dan data test sebesar 20 persen. Jumlah data yang didapat untuk data train sebesar 331 records dan untuk jumlah data untuk data testing sebesar 83 records.

```
In [52]: from sklearn.linear_model import LinearRegression
model3=LinearRegression()
model3.fit(x_train, y_train)
model3.score(x_test, y_test)
Out[52]: -0.021709161095336826
```

Gambar 4.7 Score Fitting Model Skenario Ketiga

Dari Gambar 4.12 hasil fitting model yang disajikan seperti script code gambar diatas, score yang didapatkan pada skenario ketiga yakni sebesar -0.021709161095336826.

#### 4.3.4. Memilih Skenario Splitting Data

Tabel 4.3 Skenario Splitting Data

skenario	Train	Testing	coefisien
Pertama	50	50	0.7075627489946119
Kedua	70	30	0.8041320270845731
Ketiga	80	20	-0.021709161095336826

Dilihat dari tabel 4.3 skore tertinggi yang didapat pada fitting model ada pada skenario ketiga yakni sebesar 80 persen. Maka untuk pengujian model digunakan adalah skenario kedua yakni splitting data 70 persen untuk data train dan 30 persen untuk data testing.

#### 4.4. Pengujian Model

##### 4.4.1. Pengujian *Matrix*

```
In [63]: from sklearn import metrics
MAE_linear=metrics.mean_absolute_error(y_test , y_pred)
MSE_linear=metrics.mean_squared_error(y_test , y_pred)
RMSE_linear=np.sqrt(MSE_linear)
pd.DataFrame([MAE_linear,MSE_linear,RMSE_linear], index=['MAE_linear','MSE_linear','RMSE_linear'],columns=['Quantity'])
```

```
Out[63]:
```

	Quantity
MAE_linear	3.373572
MSE_linear	47.377410
RMSE_linear	6.883125

Gambar 4.8 Pengujian Matrix

1. Mean Absolute Error (MAE) adalah dua diantara banyak metode untuk mengukur tingkat keakuratan suatu model peramalan. Nilai MAE merepresentasikan rata – rata kesalahan (error) absolut antara hasil peramalan dengan nilai sebenarnya. Nilai yang didapatkan dengan pengujian dengan MAE sebesar 3 persen.
2. Mean Squared Error (MSE) adalah Rata-rata Kesalahan kuadrat diantara nilai aktual dan nilai peramalan, dari hasil pengujian dengan menggunakan MSE nilai yang didapat 47 persen,
3. Root Mean Squared Error (RMSE) merupakan salah satu cara untuk mengevaluasi model regresi linear dengan mengukur tingkat akurasi hasil perkiraan suatu model. Dari pengujian model dengan menggunakan teknik skor yang didapat sebesar 6 persen.

#### 4.4.2. Melakukan Prediksi

Setelah melakukan pengujian pada model yang telah dibuat dengan menggunakan matrix MSE, maka dilakukan testing prediksi terhadap model yang telah dibangun. berikut tampilan source code untuk melakukan prediksi penjualan.

```
data=[
    [70000,700000, 7500000, 1, '1', '5'],
    [150000,17500000, 1750000, 29, '2', '5'],
    [150000,17500000, 1750000, 29, '3', '5'],
    [150000,17500000, 1750000, 29, '4', '5'],
    [150000,17500000, 1750000, 29, '5', '5'],
    [150000,17500000, 1750000, 29, '6', '5'],
    [150000,17500000, 1750000, 29, '7', '5'],
    [150000,17500000, 1750000, 29, '8', '5'],
    [150000,17500000, 1750000, 29, '9', '5'],
    [150000,17500000, 1750000, 29, '10', '5'],
    [150000,17500000, 1750000, 29, '11', '5'],
    [150000,17500000, 1750000, 29, '12', '5']
]
dt_rang=pd.DataFrame(data, index=['Rang', 'Rang', 'Rang', 'Rang', 'Rang', 'Rang'])
prediksi_rang=model.predict(dt_rang)
print(prediksi_rang)
```

```
[12.66063688 26.95128539 26.80375105 26.65621671 26.50868237 26.36114804
 26.2136137 26.06607936 25.91854502 25.77101068 25.62347634 25.47594201]
```

Gambar 4.9 Source Code Data Prediksi Rang

Prediksi dilakukan dengan menginputkan data penjualan produk Rang pada setiap bulan. Data-data yang diinputkan seperti data harga beli, harga jual, total harga, bulan penjualan serta

indeks dari barang. Pada code gambar diatas data yang akan diprediksi ialah produk jenis rang dan hasil dari prediksi disajikan dalam bentuk gambar dataframe.

	hbeli	hjual	totalharga	Bulan	prediksi
0	30000	350000	350000	Januari	12.660837
1	30000	350000	350000	Pebruari	26.951205
2	30000	350000	350000	Maret	26.803751
3	30000	350000	350000	April	26.656217
4	30000	350000	350000	Mei	26.508682
5	30000	350000	350000	Juni	26.361141
6	30000	350000	350000	Juli	26.213614
7	30000	350000	350000	Agustus	26.066079
8	30000	350000	350000	September	25.918545
9	30000	350000	350000	Oktober	25.771011
10	30000	350000	350000	November	25.623476
11	30000	350000	350000	Desember	25.475942

Gambar 4.10 Dataframe Hasil Prediksi Penjualan Produk Rang

Dari gambar 4.18 menunjukkan hasil prediksi dimana pada bulan januari jumla penjualan produk Rang sebanyak 12 lebih produk yang akan terjual, sedang dari bulan pebruari sampai bulan agustus sebanyak 26 lebih produk rang yang akan terjual serta dari bulan september sampai bulan desember sebanyak 25 lebih produk yang akan terjual.

## 5. KESIMPULAN DAN SARAN

### 5.1. Kesimpulan

1. Pada pemodelan dengan data train 70 persen dan data testing 30 nilai score yang didapat sebesar 80 persen. Sedangkan nilai coefisean rata-rata sebanyak 2 dan nilai intersept 7, 62.
2. Nilai hasil uji model regresi dengan menggunakan matrik mean Mean Square Error (MSE) nilai error yang didapatkan terlalu tinggi yakni sebesar 47.377, dan nilai Root Mean Square Error (RMSE) sebesar 6.883125, Sedang skor MAE sebesar 3.373572.
3. Dari hasil predik pada salah satu produk penjualan didapatkan pada bulan januari jumla penjualan produk Rang sebanyak 12 lebih produk yang akan terjual, sedang dari bulan pebruari sampai bulan agustus sebanyak 26 lebih produk rang yang akan terjual serta dari bulan september sampai bulan desember sebanyak 25 lebih produk yang akan terjual.
4. melihat dari hasil skore prediksi dan pengujian dengan menggunakan matrik, maka pemodelan dengan menggunakan algoritma regresi untuk prediksi penjualan produk UMKM masih kurang bagus dilihat dari matrik errornya lebih dari 47 persen

### 5.2. .Saran

Saran untuk penelitian selanjutnya adalah melakukan pemodelan data dengan algoritma regresi lainnya seperti polynomial regersi dan atau dengan algorit knn regresi. Dan menggunakan data yang lebih banyak..

## DAFTAR PUSTAKA

- [1] M. Idris, "pengertian UMKM," Kompas.com, 2021.
- [2] S. Suriyati, "Desain Perancangan E-Commerce Kain Tenun Lombok Desa Sukarara Lombok Tengah," To Maega | J. Pengabd. Masy., vol. 2, no. 2, p. 36, 2019, doi: 10.35914/tomaega.v2i2.241.
- [3] aris kurniawan, "materi data mining," pendidikan.co.id, 2021.

- [4] A. A. Muhartini, O. Sahroni, S. D. Rahmawati, T. Febrianti, and I. Mahuda, "Analisis Peramalan Jumlah Penerimaan Mahasiswa Baru Dengan Menggunakan Metode Regresi Linear Sederhana," *J. Bayesian J. Ilm. Stat. dan Ekon.*, vol. 1, no. 1, pp. 17–23, 2021.
- [5] D. A. Trianggana, "a Peramalan Jumlah Siswa-Siswi Melalui Pendekatan Metode Regresi Linear," *J. Media Infotama*, vol. 16, no. 2, pp. 115–120, 2020, doi: 10.37676/jmi.v16i2.1149.
- [6] Hendra and R. Harman, "Penerapan Data Mining Untuk Prediksi Penjualan Readymix Menggunakan Metode Algoritma C4.5 Pada Pt Remicon Widyaprima," *J. Comasie*, vol. 3, no. 3, pp. 84–93, 2020.
- [7] I. M. Kamal, T. H. P, and R. Ilyas, "Prediksi Penjualan Buku Menggunakan Data Mining," *Semin. Nas. Teknol. Inf. dan Multimed.*, pp. 49–54, 2017.
- [8] J. Eska, "Penerapan Data Mining Untuk Prediksi Penjualan Wallpaper Menggunakan Algoritma C4.5," vol. 2, 2018, doi: 10.31227/osf.io/x6svc.
- [9] Wahyuni dkk. (2019) "Sistem Peramalan Penjualan Perumahan DI Kabupaten Lamongan DEngan Menggunakan Metode Regresi Linear Berganda" vol.3, 2019 published September 21,2019.
- [10] Ikhsan dkk. (2018) "Regresi Linear untuk Prediksi Permintaan Sperapart Sepeda Motor" vol.3 No.2 information system 2018.