

## LINEAR REGRESSION FOR DISCOUNTING PRESENTATION RECOMMENDATIONS (Kaggle Dataset)

Nur Fahmi Maulidi<sup>1</sup>, Nur Nafiyah<sup>2</sup>

<sup>1</sup>Mahasiswa Teknik Informatika Universitas Islam Lamongan

Jalan Veteran 53A Lamongan

<sup>2</sup>Dosen Teknik Informatika Universitas Islam Lamongan

Jalan Veteran 53A Lamongan, e-mail: [mynaff@unisla.ac.id](mailto:mynaff@unisla.ac.id)

---

### ARTICLE INFO

Article history:

Received : 31 – Maret - 2022

Received in revised form : 6 – April - 2022

Accepted : 9 – Juni - 2022

Available online : 1 – September - 2022

---

### ABSTRACT

In the business of selling goods, there must be goods that do not sell well and sell well. How to make unsold items sell well by giving customers a discount or discount strategy. The goal is to provide discounted prices to attract customers' attention and increase sales turnover. Prediction to give the right discount presentation is needed in the discount strategy. How to determine discount prediction using linear regression method, looking for line equations by training data taken from Kaggle.com. The data were trained to find the constants and coefficients of the independent variables. The resulting line equation will be tested to determine the discount prediction that will be given and calculate the error value. The data used is painting sales data taken from Kaggle.com, a total dataset of 1525, and a prediction calculation is made with 225 lines of test data. The input variable used is the price of the painting, and the output variable is the discount presentation. The process of evaluating the prediction results by calculating the MAE value, the difference between the actual data and the predicted data, the MAE value from the trial is 0.108.

**Keywords:** linear regression, prediction, discount presentation.

---

### Abstrak

Dalam usaha penjualan barang pasti ada barang yang tidak laku terjual secara laris dan terjual laris. Cara membuat barang yang tidak terjual laris menjadi laris dengan strategi memberikan potongan harga atau diskon kepada pelanggan. Tujuan memberikan potongan harga untuk menarik perhatian pelanggan dan menaikkan omset penjualan. Prediksi untuk memberikan presentasi diskon yang tepat sangat dibutuhkan dalam strategi pemberian diskon. Cara menentukan prediksi diskon dengan menggunakan metode Regresi Linear, mencari persamaan garis dengan mentraining data yang diambil dari Kaggle.com. Data ditraining untuk mencari nilai konstanta dan koefisien dari variabel independen. Persamaan garis yang dihasilkan akan dilakukan testing untuk mengetahui prediksi diskon yang akan diberikan serta menghitung nilai errornya. Data yang digunakan adalah data penjualan lukisan yang diambil dari Kaggle.com, total dataset 1525 dan dilakukan perhitungan prediksi dengan data tes sebanyak 225 baris. Variabel input yang digunakan adalah

harga lukisan, dan variabel output adalah presentasi diskon. Proses evaluasi hasil prediksi dengan menghitung nilai MAE, nilai selisih antara data aktual dengan data prediksi, nilai MAE dari ujicoba adalah 0,108.

**Kata Kunci:** Regresi Linear, prediksi, presentasi diskon.

## 1. PENDAHULUAN

Memberikan pelayanan kepada pelanggan merupakan cara meningkatkan omset penjualan. Penjualan bisa meningkat jika konsumen diberi iming-iming diskon atau potongan harga. Cara yang dilakukan perusahaan salah satunya dengan memberikan potongan harga atau diskon pada konsumen pada produk-produk tertentu [1].

Dalam dunia industri atau perusahaan harus selalu mengembangkan strategi penjualan. Tujuannya agar meningkatkan omset laba atau keuntungan dari perusahaan. Salah satu cara untuk meningkatkan laba adalah membuat suatu sistem penentuan diskon atau potongan harga kepada pelanggan atau konsumen [2].

Penentuan diskon selain untuk meningkatkan laba juga sebagai sarana agar barang-barang yang tidak laku terjual bisa laris terjual. Tujuan memberikan diskon selain menggait pelanggan, meningkatkan omset/laba, bisa juga agar barang-barang yang stoknya masih banyak atau tidak laku bisa laris terbeli [3]. Tujuan dari memberikan potongan harga pada pelanggan agar mengurangi kerugian dari biaya produksi barang tersebut. Salah satu metode yang dapat memberikan prediksi presentasi diskon yang diberikan ke pelanggan adalah Regresi Linear. Metode Regresi Linear dapat digunakan untuk memprediksi harga jual mobil bekas [4], selain itu Regresi Linear berganda juga dapat digunakan untuk memprediksi produksi hasil panen padi [5], Regresi Linear sangat baik digunakan memprediksi nilai CPO [6]. Hasil evaluasi metode Regresi Linear dengan menghitung MAE mempunyai nilai yang rendah, serta selisih errornya antara data aktual dengan hasil prediksi kecil [7].

Dari penelitian sebelumnya, peneliti bermaksud membuat perhitungan presentasi diskon pada lukisan dengan metode Regresi Linear. Tujuan penelitian ini agar dapat meningkatkan penjualan, serta mengurangi kerugian.

## 2. TINJAUAN PUSTAKA

Pada penelitian [8] terkait menentukan variabel bebas dan variabel tidak bebas, nilai konstanta, koefisien regresi, serta menginterpretasikan koefisien regresi, nilai koefisien determinasi, koefisien korelasi berganda. Juga menghitung *standart error estimate* dan nilai parsialnya untuk mengetahui kisaran biaya pencetakan sebuah spanduk. Hasil akhirnya bahwa metode regresi linier berganda cocok digunakan estimasi biaya pencetakan spanduk. Penelitian lain oleh [9] metode Regresi Linier Sederhana dapat digunakan mencari korelasi pengaruh biaya diskon terhadap banyaknya penjualan, dengan menghitung nilai koefisien regresi.

Penelitian terkait pengaruh promosi dengan diskon dilakukan oleh [10]. Metode yang digunakan untuk melihat hubungan antara lokasi, harga, promosi dan kualitas pelayanan sebagai variabel independen (x) dan Citra merek sebagai variabel dependen (y) adalah Regresi Linear berganda. Hasil akhirnya dilakukan pengujian untuk melihat nilai koefisien signifikan dan determinasi.

Metode Regresi Linear memiliki 3 kegunaan yaitu untuk menggambarkan sebuah permasalahan yang diteliti, kedua untuk tujuan kontrol, ketiga untuk memprediksi. Rumus Regresi Linear sederhana merupakan Persamaan 1.

$$\hat{y} = \beta_0 + \beta_1 x_1 \quad (1)$$

Keterangan Persamaan 1  $\hat{y}$  adalah persamaan garis,  $\beta_0$  adalah nilai konstanta,  $\beta_1$  adalah koefisien dari variabel  $x_1$  atau variabel independen dan  $x_1$  adalah variabel x ke-i. Cara mencari  $\beta_0$  atau nilai konstanta dapat menghitung dengan Persamaan 2.

$$\beta_0 = \bar{y} - \beta_1 \bar{x} \quad (2)$$

Keterangan Persamaan 2  $\bar{y}$  merupakan rata-rata y dan  $\bar{x}$  adalah rata-rata x. untuk mencari  $\beta_1$  atau nilai koefisien variabel  $x_1$  dapat dilakukan dengan Persamaan 3.

$$\beta_1 = \frac{\sum_{i=1}^n y_i x_i - \frac{(\sum_{i=1}^n y_i)(\sum_{i=1}^n x_i)}{n}}{\sum_{i=1}^n (x_i)^2 - \frac{(\sum_{i=1}^n x_i)^2}{n}} \quad (3)$$

Keterangan Persamaan 3  $y_i$  adalah variabel  $y$  ke- $i$  dan  $x_i$  adalah variabel  $x$  ke- $i$  dan  $n$  merupakan banyak data yang digunakan. Cara menghitung evaluasi metode Regresi Linear dengan MAE seperti Persamaan 4.

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (4)$$

Keterangan Persamaan 4  $y$  adalah data aktual, dan  $\hat{y}$  adalah hasil prediksi.

### 3. METODOLOGI PENELITIAN

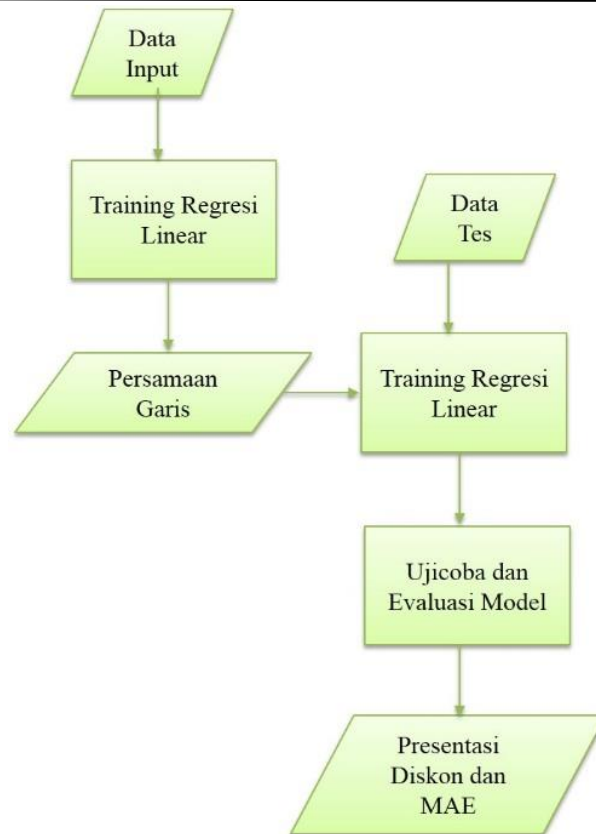
Data yang digunakan dalam penelitian ini diambil dari website kaggle.com. Total data sebanyak 1525 baris, data terkait seni lukisan, seni patung, dan seni lainnya. Data dibagi menjadi data training sebanyak 1300 dan data tes sebanyak 225 baris. Data input adalah harga seni dalam satuan harga Euro, dan data output atau variabel output adalah jumlah diskon dalam satuan persen. Tabel 1 merupakan contoh dataset yang digunakan.

Tabel 1. Contoh Dataset

Harga	diskon (%)
2,906	36
2,999	25
1,049	15
3,635	38
6,299	10
4,499	23
4,841	22
5,499	21
4,578	5
5,192	29
5,099	39
5,549	30

Tabel 1 menunjukkan bahwa harga adalah harga barang-barang seni dalam satuan Euro, dan diskon dalam satuan %.

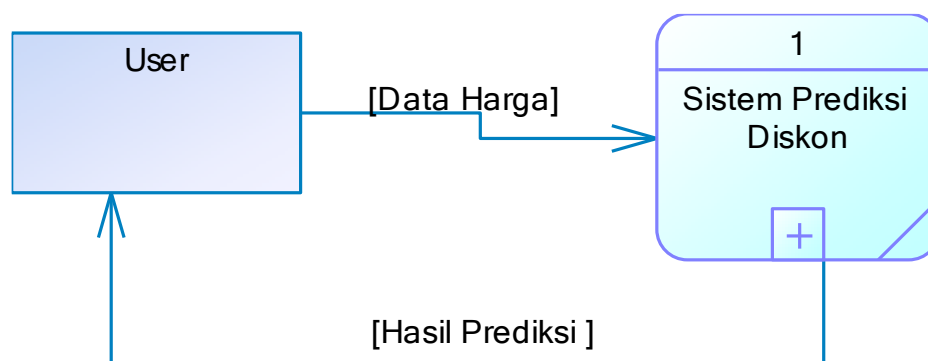
Penelitian ini akan melakukan prediksi presentasi diskon dari suatu barang dengan metode Regresi Linear. Alur dari proses prediksi seperti dalam Gambar 1.



Gambar 1. Alur Sistem Prediksi Diskon

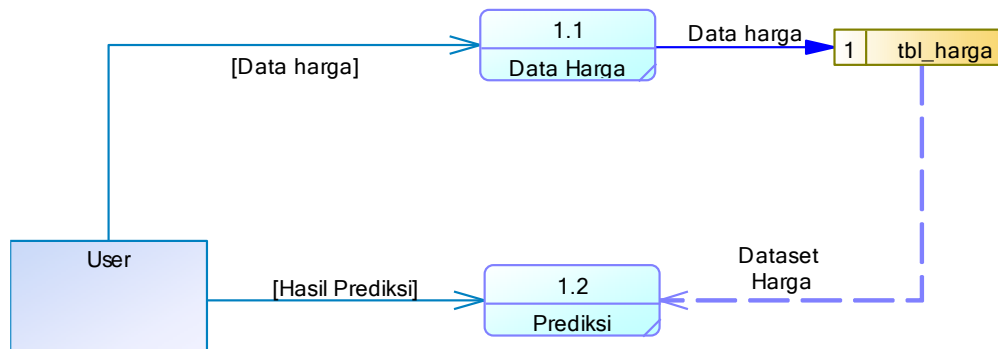
Gambar 1 data diinputkan dalam bentuk file csv dan diproses training serta testing menggunakan Jupyter Notebook dengan bahasa Python. Hasil prediksi akan dihitung nilai errornya dengan menghitung nilai MAE Persamaan 4.

Diagram Konteks ditampilkan dalam Gambar 2 yang menggambarkan alur sistem yang dibuat untuk memudahkan dalam perancangan sistem.



Gambar 2. Diagram Konteks

*Data Flow Diagram* (DFD) level 0 merupakan gambaran rancangan sistem dan alur secara lebih rinci dan dapat menjelaskan sistem secara keseluruhan.



Gambar 3. DFD Level 0

Pada Gambar 3 merupakan gambaran alur sistem prediksi presentasi diskon secara rinci dengan user menginputkan data harga berbentuk csv, kemudian dari data harga tersebut dilakukan training perhitungan menggunakan metode Regresi Linear selanjutnya user dapat melihat dan mendapatkan hasil prediksi.

#### 4. HASIL DAN PEMBAHASAN

Tabel 1 jika dihitung dan ditraining agar menghasilkan persamaan seperti dalam contoh hitungan Tabel 2.

Tabel 2 Contoh Proses Training Mendapatkan Persamaan

Harga (x)	diskon (%) (y)	x.y	$x^2$
2906	36	104616	8444836
2999	25	74975	8994001
1049	15	15735	1100401
3635	38	138130	13213225
6299	10	62990	39677401
4499	23	103477	20241001
4841	22	106502	23435281
5499	21	115479	30239001
4578	5	22890	20958084
5192	29	150568	26956864
5099	39	198861	25999801
5549	30	166470	30791401

$$\beta_1 = \frac{1260693 - \frac{(293 \times 52145)}{12}}{250051297 - \frac{2719101025}{12}} = -0.00053$$

$$\beta_0 = 24.42 - (-0.00053 \times 4345.42) = 26.73$$

$$\hat{y} = 26.73 - 0.00053 x_1$$

(5)

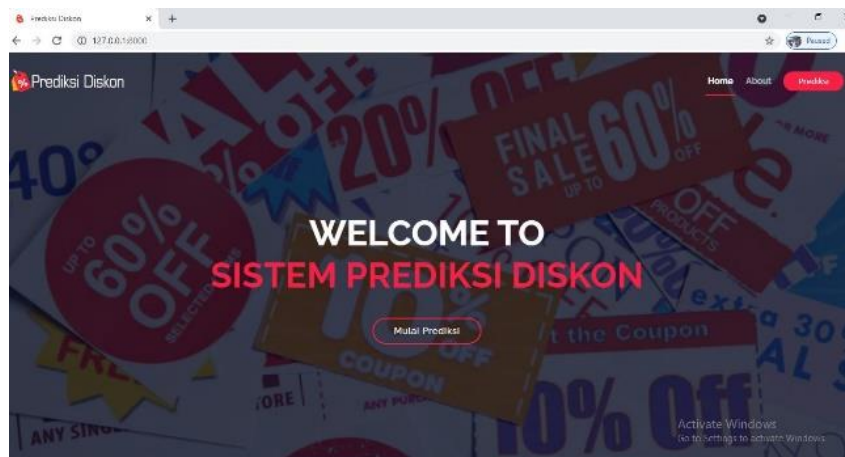
Data Tabel 1 dilakukan training seperti proses di Tabel 2 maka menghasilkan persamaan garis seperti Persamaan 5.

Data 1300 baris dilakukan training di Jupyter Notebook menghasilkan persamaan garis seperti berikut  $\hat{y} = 0,403 - 0,00001 x_1$ . Data tes 225 baris digunakan untuk evaluasi nilai MAE adalah 0,108. Hasil training Regresi Linear di Jupyter Notebook kemudian disimpan modelnya dalam bentuk pickle dan model tersebut digunakan membuat sistem web prediksi presentasi diskon. Hasil evaluasi dari model Regresi Linear untuk memprediksi presentasi diskon seperti Gambar 4. Gambar 4 warna titik merah adalah data hasil prediksi, sedangkan warna titik biru adalah data aktual.



Gambar 4. Hasil Evaluasi Prediksi

Sistem prediksi presentasi diskon ini dibangun menggunakan bahasa pemrograman python dengan bantuan Jupyter Notebook (digunakan proses training) dan tampilan interfacenya menggunakan bahasa pemrograman python, html dan css. Hasil implementasi dari perancangan sistem yang telah dibuat adalah menu yang paling awal dijalankan oleh sistem berisi tampilan halaman dengan gambar dan tombol menu-menu yang dapat diakses ketika diklik atau ditekan. Dalam halaman utama berisi menu *home*, *about* dan menu prediksi. Tampilan halaman utama seperti pada Gambar 5.



Gambar 5. Halaman Utama

Halaman Prediksi merupakan halaman untuk melakukan perhitungan dalam menentukan prediksi presentasi diskon yang akan diberikan. Gambar 6 tampilan proses prediksi.



Gambar 6. Halaman Prediksi

## 5. KESIMPULAN DAN SARAN

Data 1300 baris dilakukan training di Jupyter Notebook menghasilkan persamaan garis seperti berikut  $\hat{y} = 0,403 - 0,00001 x_1$ . Data tes 225 baris digunakan untuk evaluasi nilai MAE adalah 0,108. Hasil training Regresi Linear di Jupyter Notebook kemudian disimpan modelnya dalam bentuk pickle dan model tersebut digunakan membuat sistem web prediksi presentasi diskon.

## Ucapan Terima Kasih

Terima kasih kepada kedua orang tua dan dosen yang telah memberikan bimbingan serta arahan.

## 6. DAFTAR PUSTAKA

- [1] T. Kurnialensya and R. Abidin, "SISTEM PENDUKUNG KEPUTUSAN PELANGGAN TERBAIK DAN PEMBERIAN DISKON MENGGUNAKAN METODE SAW & TOPSIS," *J. Ilm. Elektron. DAN Komput.*, vol. 13, no. 1, pp. 18–33, 2020.
- [2] G. Sianipar and M. Marbun, "Sistem Pendukung Keputusan Penentuan Jumlah Diskon Produk Pada PT . Indomarco Prismatama Beringin 5-7 Tembung Menggunakan SAW," *J. Inform. Pelita Nusant.*, vol. 5, no. 1, pp. 1–6, 2020.
- [3] E. Sonalitha and B. Nurdewanto, "Sistem Penentuan Diskon Pada Swalayan Berbasis Jumlah Penjualan Dan Stok Barang Menggunakan Metode Fuzzy Control," *Matics*, vol. 7, no. 1, pp. 43–48, 2015, doi: 10.18860/mat.v7i1.2876.
- [4] N. Nafi'iyah, "Penerapan regresi linear dalam memprediksi harga jual mobil bekas," pp. 1–5, 2015.
- [5] E. Triyanto, H. Sismoro, and A. D. Laksito, "IMPLEMENTASI ALGORITMA REGRESI LINEAR BERGANDA UNTUK MEMPREDIKSI PRODUKSI PADI DI KABUPATEN BANTUL," *Rabit J. Teknol. dan Sist. Inf. Univrab*, 2019, doi: 10.36341/rabit.v4i2.666.
- [6] A. Fitri Boy, "Implementasi Data Mining Dalam Memprediksi Harga Crude Palm Oil (CPO) Pasar Domestik Menggunakan Algoritma Regresi Linier Berganda (Studi Kasus Dinas Perkebunan Provinsi Sumatera Utara)," *J. Sci. Soc. Res.*, vol. 4307, no. 2, pp. 78–85, 2020, [Online]. Available: <http://jurnal.goretanpena.com/index.php/JSSR>.
- [7] A. A. Suryanto, "PENERAPAN METODE MEAN ABSOLUTE ERROR (MEA) DALAM ALGORITMA REGRESI LINEAR UNTUK PREDIKSI PRODUKSI PADI," *SAINTEKBU*, 2019, doi: 10.32764/saintekbu.v1i1.298.
- [8] A. Rivandi, E. Bu'ulolo, and N. Silalahi, "PENERAPAN METODE REGRESI LINIER BERGANDA DALAM ESTIMASI BIAYA PENCETAKAN SPANDUK (STUDI KASUS : PT. HANSINDO SETIAPRATAMA) Ahmad," *Pelita Inform. Budi Darma*, vol. 18, 2019.
- [9] I. B. K. A. Wardana, "Pengaruh Biaya Promosi Terhadap Volume Penjualan Patung UD," *J. Jur. Pendidik. Ekon.*, vol. 5, no. 1, pp. 1–11, 2015.
- [10] D. R. Saputra, "Analisis Pengaruh Lokasi, Harga, Promosi, dan Kualitas Pelayanan terhadap Citra Merek," *Univ. Muhammadiyah Jember*, 2017.