



**Indonesia AI**  
AI for Everyone, AI for Indonesia

**# 2023JAGO**  
**TEKNOLOGIAI**



**AI Career Bootcamp**

# **Linear & Logistic Regression**

## Pembukaan

Guide Book ini memuat beberapa informasi-informasi utama yang akan disampaikan oleh mentor di program AI Career Bootcamp Kelas Computer Vision yang bisa dijadikan pegangan para students untuk mempersiapkan diri sebelum sesi Live Class berlangsung.

## Deskripsi

Linear Regression merupakan algoritma yang digunakan dalam aplikasi Machine Learning untuk memodelkan hubungan linear antara variabel independen dan dependen. Algoritma ini digunakan untuk memprediksi nilai numerik dari variabel dependen berdasarkan nilai numerik dari satu atau lebih variabel independen. Model Linear Regression menghasilkan persamaan garis lurus dengan koefisien yang dapat digunakan untuk memprediksi nilai variabel dependen berdasarkan nilai variabel independen. Model ini sangat berguna dalam pengambilan keputusan bisnis dan prediksi penjualan berdasarkan faktor-faktor yang mempengaruhinya.

Logistic Regression adalah algoritma yang digunakan dalam aplikasi Machine Learning untuk memodelkan hubungan antara satu atau lebih variabel independen dan variabel dependen biner. Variabel dependen dalam algoritma ini hanya memiliki dua nilai, yaitu 0 dan 1, yang merepresentasikan keberadaan atau ketiadaan suatu kejadian atau kondisi. Algoritma ini digunakan untuk memprediksi probabilitas kejadian atau kondisi berdasarkan variabel independen. Model logistic regression menghasilkan kurva sigmoid yang dapat digunakan untuk memprediksi nilai probabilitas suatu kejadian atau kondisi. Model ini sering digunakan dalam prediksi kelayakan kredit, pengambilan keputusan medis, dan penentuan keterlibatan dalam kejahatan.

Kedua algoritma ini memiliki kegunaan yang berbeda-beda dalam aplikasi Machine Learning. Linear Regression digunakan untuk memodelkan hubungan linear antara variabel independen dan dependen, sedangkan Logistic Regression digunakan untuk memodelkan hubungan antara variabel independen dan variabel dependen biner. Namun, keduanya memiliki satu kesamaan yaitu keduanya merupakan algoritma yang paling umum digunakan dalam aplikasi Machine Learning dengan metode Supervised Learning untuk memprediksi nilai variabel dependen berdasarkan nilai variabel independen.

## Contoh Kasus

Berikut merupakan contoh bagaimana algoritma Linear Regression dan Logistic Regression digunakan untuk membangun sistem cerdas. Dalam hal ini, pertama kita perlu mendefinisikan variabel independen dan dependen yang dibutuhkan.

Contoh Linear Regression:



Variabel independen :

Luas tanah

Variabel dependen :

Harga rumah

Dalam contoh ini, kita ingin mengetahui apakah ada hubungan linear antara luas tanah dengan harga rumah. Luas tanah adalah variabel independen yang digunakan untuk memprediksi harga rumah yang merupakan variabel dependen.

Contoh Logistic Regression:



Variabel independen:

Tingkat pendidikan

Variabel dependen biner:

Kemungkinan seseorang mendapatkan pekerjaan (Ya/Tidak, Yes/No)

Dalam contoh ini, kita ingin mengetahui apakah ada hubungan antara tingkat pendidikan seseorang dengan kemungkinan mendapatkan pekerjaan. Tingkat pendidikan adalah variabel independen yang digunakan untuk memprediksi kemungkinan seseorang mendapatkan pekerjaan, yang merupakan variabel dependen yang hanya memiliki dua nilai, yaitu 0 dan 1, yang merepresentasikan

## Kelebihan & Kekurangan

Linear Regression dan Logistic Regression memiliki kelebihan dan kekurangan masing-masing. Berikut ini adalah penjelasannya:

Kelebihan dari Linear Regression adalah sebagai berikut:

- Mudah dipahami dan diimplementasikan
- Menghasilkan hasil yang mudah diinterpretasikan
- Efisien dalam memproses data dengan fitur linear
- Dapat digunakan dalam prediksi variabel numerik

Namun, Linear Regression juga memiliki beberapa kekurangan, di antaranya:

- Hanya dapat digunakan untuk data yang memiliki hubungan linear
- Tidak dapat digunakan untuk memprediksi variabel biner atau kategori

Sedangkan, Kelebihan dari Logistic Regression sebagai berikut:

- Mudah dipahami dan diimplementasikan
- Dapat digunakan untuk memprediksi variabel biner atau kategori
- Menghasilkan hasil yang mudah diinterpretasikan
- Efisien dalam memproses data dengan fitur kategori

Adapun kekurangannya yaitu sebagai berikut:

- Tidak efektif untuk memproses data dengan fitur yang kompleks atau tidak linear
- Tidak dapat digunakan untuk memprediksi variabel numerik

Kesimpulannya, kedua algoritma ini memiliki kelebihan dan kekurangan masing-masing, tergantung pada jenis data dan tujuan analisis yang ingin dicapai. Oleh karena itu, pemilihan algoritma yang tepat harus didasarkan pada jenis data yang digunakan dan tujuan analisis yang ingin dicapai.