



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

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

**AI Career Bootcamp**

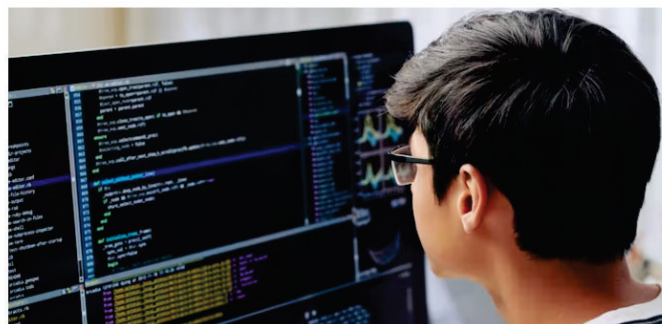
# **Deep Dive into Deep Learning (Backward)**

## 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

Backpropagation adalah teknik optimisasi yang digunakan di algoritma Deep Learning untuk memperbaiki bobot dan bias pada jaringan neural network. Tujuannya adalah untuk menyesuaikan parameter-parameter jaringan agar dapat menghasilkan output yang lebih akurat dan mendekati nilai target (berdasarkan data).



Proses backpropagation dimulai dengan menghitung nilai gradien dari Loss function terhadap parameter-parameter jaringan, yaitu bobot dan bias. Gradien ini mengindikasikan seberapa besar pengaruh perubahan setiap parameter terhadap peningkatan atau penurunan Loss Function (konsep kalkulus sebagai ilmu yang mempelajari perubahan). Kemudian, gradien tersebut dikalikan dengan faktor pembelajaran (Learning Rate) dan digunakan untuk memperbarui bobot dan bias pada algoritma neural network.

## Loss Fuction

Loss Function adalah fungsi yang digunakan untuk mengukur seberapa baik jaringan neural mampu memprediksi hasil yang diharapkan. Tujuannya adalah untuk meminimalkan selisih antara output yang dihasilkan oleh neural network dan target atau label yang diinginkan. Pada dasarnya, Loss Function merupakan metrik yang digunakan untuk mengukur seberapa akurat jaringan neural dalam melakukan prediksi.

Terdapat beberapa jenis Loss Function yang dapat digunakan pada deep learning, tergantung pada jenis tugas yang ingin diselesaikan. Contohnya, jika tugasnya adalah klasifikasi biner (ya atau tidak), Loss Function yang umum digunakan adalah binary cross-entropy. Jika tugasnya adalah klasifikasi multi kelas (lebih dari dua kelas), Loss Function yang umum digunakan adalah categorical cross-entropy. Sedangkan jika tugasnya adalah regresi (prediksi nilai yang kontinu), Loss Function yang umum digunakan adalah mean squared error (MSE).

Proses training neural network melibatkan pengoptimalan parameter-parameter (bobot/weight dan bias) untuk meminimalkan Loss Function. Proses ini dilakukan dengan menggunakan teknik backpropagation, di mana gradien pada Loss Function akan dihitung dan digunakan untuk memperbarui bobot dan bias pada neural network tersebut. Semakin kecil nilai Loss Function, semakin akurat neural network dalam melakukan prediksi.

Contoh Loss Function untuk output dari Neural Network berupa nilai regresi:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

Sederhananya, formula diatas mencoba menghitung jarak seberapa jauh nilai Label (Y) terhadap nilai hasil prediksi Neural Network ( $\hat{Y}$ ).

Pemilihan jenis Loss Function yang tepat sangat penting dalam mendapatkan hasil yang optimal pada deep learning. Jika jenis Loss Function yang digunakan tidak cocok dengan tugas yang diberikan, maka hasil prediksi jaringan dapat menjadi tidak akurat atau bahkan tidak konvergen. Oleh karena itu, pemilihan Loss Function harus dilakukan dengan hati-hati dan disesuaikan dengan jenis tugas yang ingin diselesaikan.

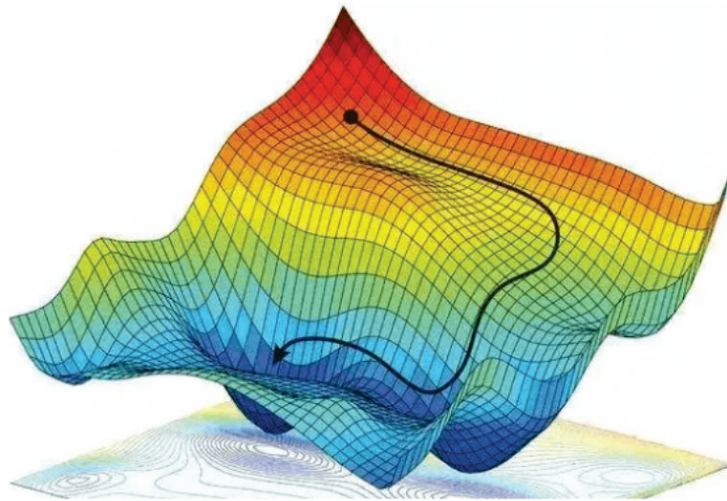
## Gradient Descent

Gradient Descent adalah salah satu teknik optimisasi yang digunakan untuk menemukan nilai minimum sebuah fungsi. Teknik ini seringkali digunakan dalam Machine Learning untuk mengoptimalkan model dengan cara meminimalkan nilai loss yang dihitung menggunakan Loss Function.

Cara kerja Gradient Descent adalah dengan menghitung gradien atau turunan parsial dari fungsi loss terhadap setiap parameter (bobot/weight) model. Kemudian, nilai parameter tersebut diubah dengan arah yang berlawanan dengan gradien sehingga fungsi loss dapat menurun (konsep kalkulus sebagai ilmu yang mempelajari perubahan). Proses ini diulang-ulang hingga konvergen pada nilai minimum fungsi loss atau sejumlah iterasi yang ditentukan sebelumnya.



Ilustrasi bagaimana sebuah proses Gradient Descent bekerja:



Tanda panah hitam adalah proses perubahan nilai loss (titik hitam) dari posisi tinggi menuju posisi rendah (optimum).

Terdapat beberapa varian dari Gradient Descent seperti Batch, Stochastic, dan Mini-batch. Batch Gradient Descent akan menghitung gradien terhadap seluruh data training sekaligus, sementara Stochastic Gradient Descent hanya menghitung gradien terhadap satu data training pada setiap iterasi. Sedangkan Mini-batch Gradient Descent akan menghitung gradien terhadap sekelompok data training dengan ukuran tertentu pada setiap iterasi. Pemilihan varian Gradient Descent yang tepat akan mempengaruhi kecepatan konvergensi dan kualitas model yang dihasilkan.