**Algoritma machine learning** adalah ilmu yang mempelajari mengenai algoritma dan metode statistik yang digunakan dalam sistem komputer tanpa menerapkan instruksi tertentu. Machine learning merupakan bagian dari Artificial Intelligence (AI) dimana mesin mempelajari data di masa lalu yang digabungkan dan diolah dengan algoritma yang sesuai. Data yang dilatih di machine learning disebut training data. Semakin banyak data yang dilatih pada machine learning, maka semakin bagus algoritmanya dan output yang diberikan. Artikel kali ini akan dibahas serba-serbi algoritma machine learning. Seperti apa sih kecerdasan buatan yang populer ini? Penasaran? Simak dibawah ini, yuk!

**1. Macam-Macam Algoritma Machine Learning**

Dalam mempelajari algoritma machine learning, ada beberapa tipe machine learning yang wajib dipahami yaitu sebagai berikut:

* Supervised learning adalah algoritma machine learning yang paling populer dimana data yang digunakan dalam pembelajarannya adalah data yang berlabel. Supervised learning digunakan untuk memprediksi yang akan terjadi di masa mendatang menggunakan data lampau yang tersedia. Beberapa algoritma yang termasuk dalam supervised learning yaitu Logistic Regression, Decision Tree, Random Forest, k-Nearest Neighbor (KNN), dan Naive Bayes.
* Unsupervised learning merupakan kebalikan dari supervised learning dimana data yang digunakan dalam pembelajarannya adalah data yang tidak berlabel. Algoritma ini mengidentifikasi data berdasarkan fitur tertentu seperti kepadatan dan strukturnya. Unsupervised learning dapat menemukan pola tersembunyi pada kumpulan data tersebut. Beberapa algoritma yang bisa digunakan dalam unsupervised learning yaitu K-Means Clustering dan Hierarchical Clustering.
* Reinforcement learning adalah algoritma yang mempelajari data dengan sendirinya berdasarkan pada lingkungan yang mempengaruhinya untuk mencapai suatu tujuan tertentu. Algoritma ini dapat digunakan dalam proses pengambilan keputusan secara spesifik. Selain itu juga dapat digunakan untuk memecahkan masalah dalam bisnis.

## **Linear Regression**

## Apa itu Regression?

Sebelum melangkah lebih jauh, apakah kamu sudah tahu apa itu pengertian dari regression? Dalam ilmu statistik, teknik untuk mencari perkiraan hubungan antara variabel-variabel numerik ini disebut dengan metode **Regression Analysis**. Lebih simpelnya, regresi merupakan salah satu teknik untuk meramalkan data dimasa yang akan datang.

## Apa itu Linear Regression

**Linear Regression** (Regresi Linear) adalah suatu regresi linear yang digunakan untuk mengestimasi atau memprediksi hubungan antara dua variabel dalam penelitian kuantitatif. Dimana regresi linear ini mampu membuat satu asumsi tambahan yang mengkorelasikan antara variabel independen dan dependen melalui garis yang paling sesuai dari titik data garis lurus. Artinya, bukan kurva atau semacam faktor penelompokkan.

Meski demikian, regresi linear memiliki keterbatasan, karena dalam data terbaik pun tidak menceritakan kisah yang lengkap. Analisis regresi biasanya digunakan dalam penelitian untuk menetapkan bahwa ada korelasi antar variabel.

Namun, korelasi yang tidak sama sebab akibat adanya hubungan antara dua variabel tidak berarti yang satu menyebabkan yang lainnya terjadi. Bahkan garis dalam regresi linear sederhana yang cocok dengan titik data mungkin tidak menjamin hubungan sebab-akibat.

**Menurut Sugiyono (2011)**, regresi linear sederhana adalah regresi linear yang didasarkan pada hubungan fungsional ataupun kausal satu variabel independen dengan satu variabel dependen.

### Kelebihan dari linear regression

Ketika hubungan antara variabel independen dan dependen memiliki hubungan linear, algoritma ini adalah yang terbaik untuk digunakan, karena ini adalah yang paling kompleks dibanding algoritma lain yang juga menemukan hubungan antara variabel independen dan dependen.

Metode ini mampu digunakan untuk memprediksi nilai yang ada pada masa depan. Hal ini sejalan dengan fungsi dari analisis regresi yang dapat digunakan untuk peramalan dan prediksi.

### Kekurangan dari linear regression

Pada kenyataannya, dalam data real, jarang masalah didunia yang menunjukkan hubungan yang jelas antara variabel dependen dan independen. Hal ini dapat membuat model yang tidak cukup bagus, disebabkan karena kesalahan dalam memilih variabel yang digunakan untuk analisis.