

PERTEMUAN 4

Sains data Dan Machine Learning

Hubungan data science dan Machine learning

Analisis data dan proses bisnis dilakukan oleh datascience

Pemodelan menggunakan learning machine

Kategori Machine Learning

Supervised Learning

Unsupervised Learning

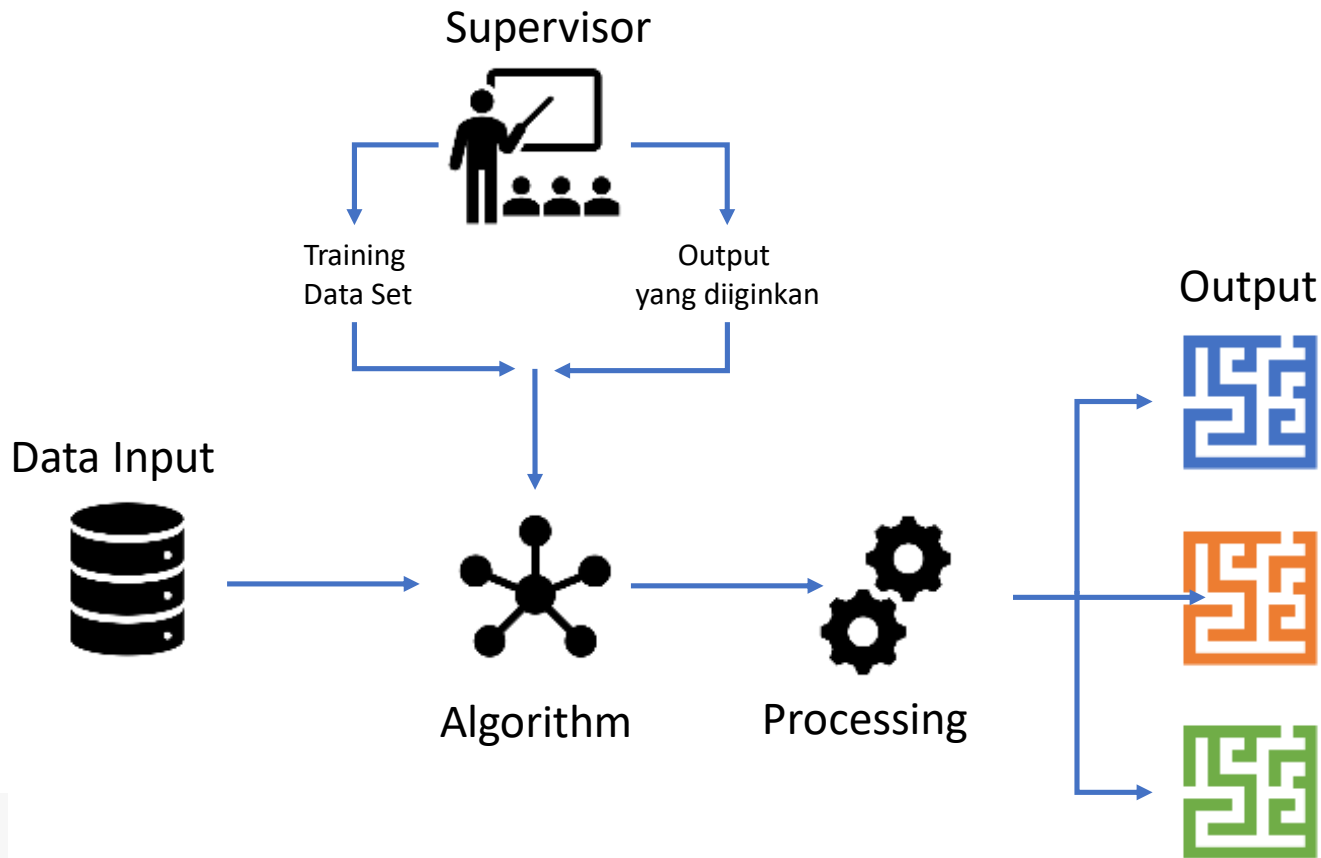
Reinforcement Learning

Supervised Learning

Supervised Learning adalah subkategori pembelajaran mesin (Machine Learning) dan kecerdasan buatan (AI).

Model ini ditentukan oleh penggunaan kumpulan data berlabel untuk melatih algoritma yang mengklasifikasikan data atau memprediksi hasil secara akurat.

Konsep Kerja Supervised Learning



Sumber : medium.com

Konsep Kerja Supervised Learning (lanjutan)

- Saat data input dimasukkan ke dalam model, ia menyesuaikan bobotnya hingga model dipasang dengan tepat, yang terjadi sebagai bagian dari proses validasi silang

Cara Kerja Supervised Learning

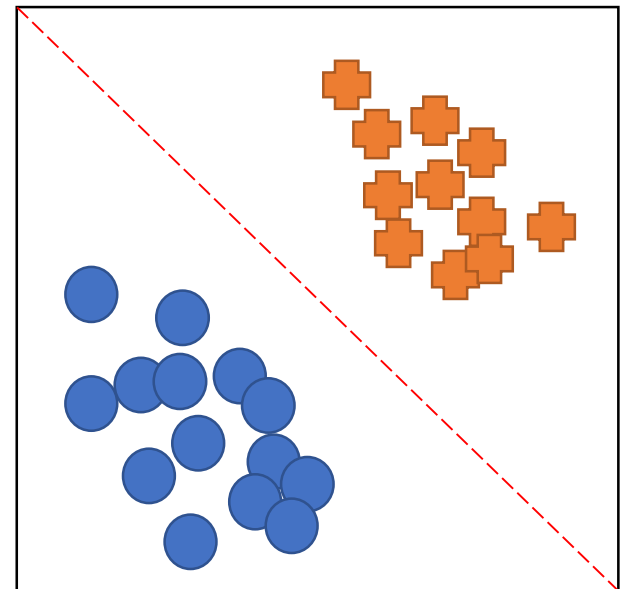
- *Supervised Learning* menggunakan satu set data training untuk melakukan learning model, guna menghasilkan output yang diinginkan.
- Dataset pada data *training* mencakup input dan output yang benar, yang memungkinkan model untuk belajar dari waktu ke waktu.
- Algoritma nantinya mengukur akurasi melalui fungsi kerugian (*Loss Function*), lalu menyesuaikan hingga kesalahan cukup diminimalkan.

Jenis supervised learning

Supervised Learning dapat dibagi kedalam 2 jenis berdasarkan cara pemecahan masalahnya.

Klasifikasi (*Classification*)

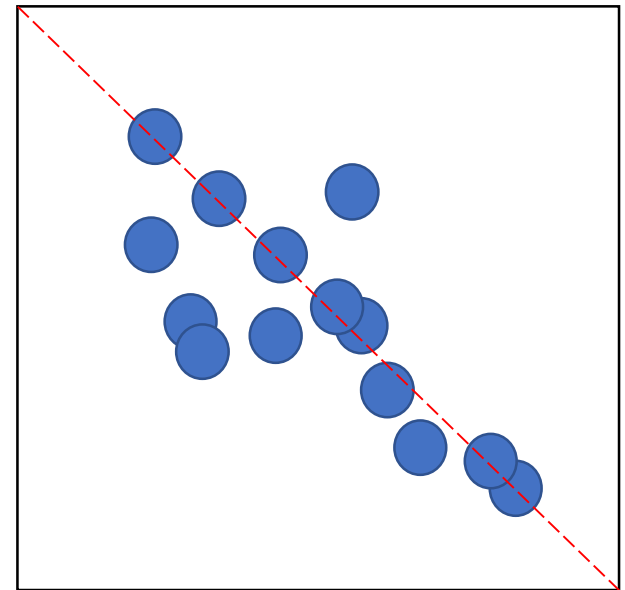
Menggunakan algoritma, untuk menetapkan data uji secara akurat menetapkan ke dalam kategori tertentu. Jenis ini mengenali entitas tertentu dalam kumpulan data dan mencoba menarik beberapa kesimpulan tentang bagaimana entitas tersebut harus diberi label atau didefinisikan. Algoritma umum yang digunakan, SVM, Decision Tree, CNN, linear classifier dll



Jenis supervised learning (lanjutan)

Regresi (*Regression*)

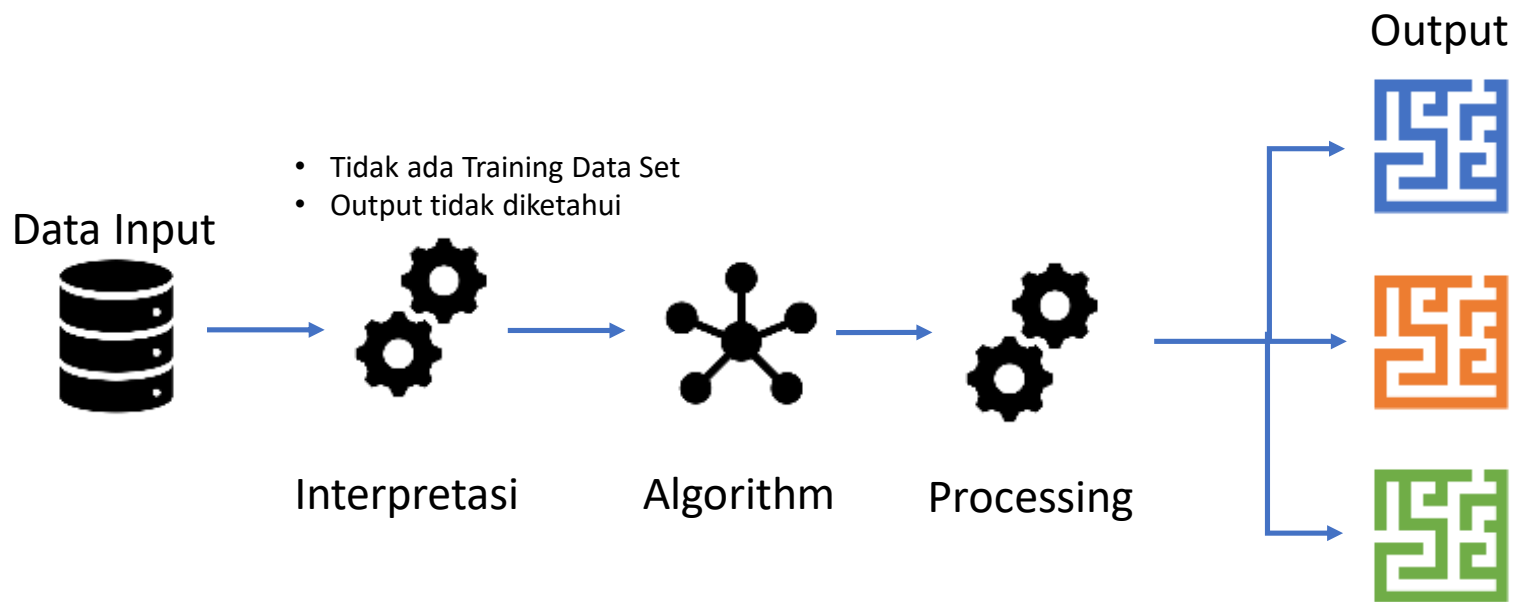
digunakan untuk mengetahui hubungan antara variabel dependen dan variabel bebas. Jenis ini biasanya digunakan untuk membuat proyeksi, seperti untuk pendapatan penjualan untuk bisnis tertentu. Regresi yang paling umum digunakan saat ini adalah : Regresi Linier, Polinomial dan logistic regression



Model *UnSupervised Learning*/Descriptive

- Unsupervised Learning menggunakan algoritma pembelajaran mesin (machine learning) untuk menganalisis dan mengelompokkan kumpulan data yang tidak berlabel.
- Algoritma ini menemukan pola tersembunyi atau pengelompokan data tanpa perlu campur tangan manusia, berbanding terbalik dengan Supervised Learning yang erat dengan campur tangan manusia.
- Kemampuan Unsupervised Learning menemukan persamaan dan perbedaan informasi menjadikannya solusi ideal untuk analisis data eksplorasi, strategi penjualan silang, segmentasi pelanggan, dan pengenalan gambar
- Contoh algoritma yang digunakan di model ini: K-Means Clustering Algorithm.

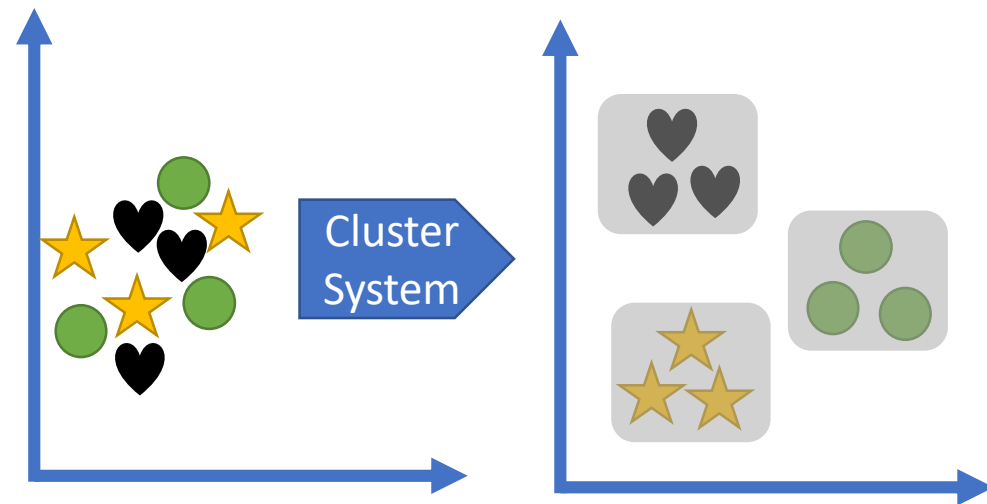
Konsep Kerja unsupervised Learning



Jenis-jenis Unsupervised Learning

Clustering

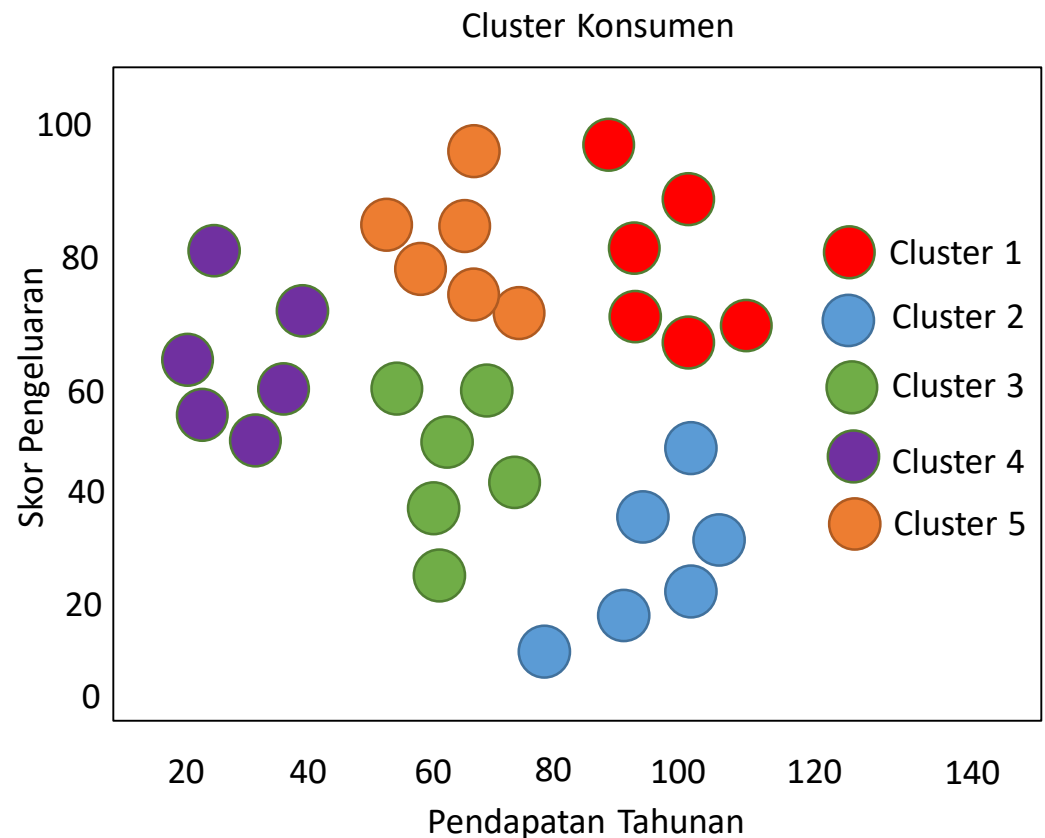
Clustering adalah teknik data mining yang mengelompokkan data yang tidak berlabel berdasarkan persamaan atau perbedaannya. Algoritma pengelompokan digunakan untuk memproses objek data mentah yang tidak diklasifikasikan ke dalam kelompok yang diwakili oleh struktur atau pola dalam informasi.



Jenis-jenis Unsupervised Learning

Hierarchical Clustering

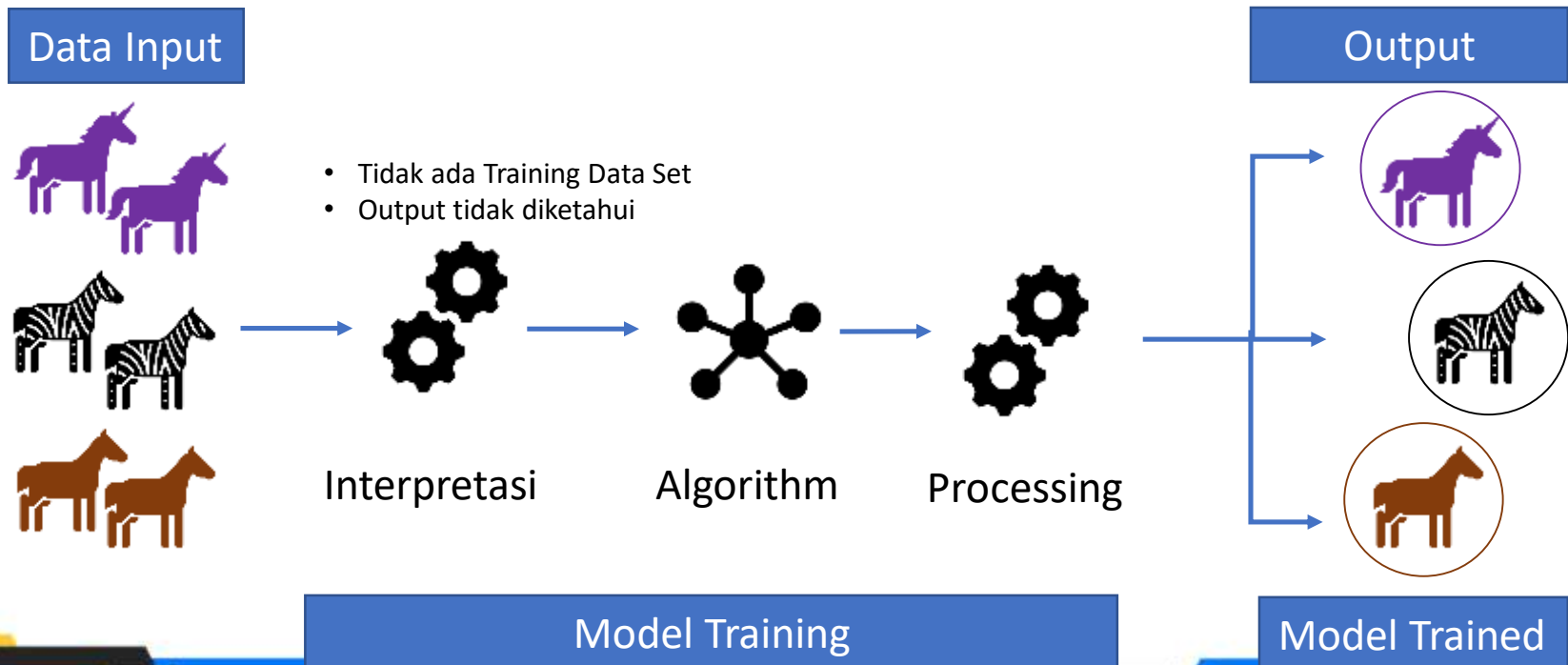
Hierarchical Clustering adalah algoritma pengelompokan tanpa pengawasan yang dapat dikategorikan dalam dua cara, yakni bisa menjadi aglomerat atau memecah belah.



Jenis-jenis Unsupervised Learning

Probabilistik Clustering

Model probabilistik adalah teknik Unsupervised Learning yang memecahkan estimasi kepadatan atau masalah pengelompokan lunak. Dalam pengelompokan probabilistik, titik data dikelompokkan berdasarkan kemungkinan bahwa mereka termasuk dalam distribusi tertentu. Model Campuran Gaussian (GMM) adalah salah satu metode pengelompokan probabilistik yang paling umum digunakan

























Jenis-jenis Unsupervised Learning

Association Rule

Association Rules adalah metode berbasis aturan untuk menemukan hubungan antara variabel dalam kumpulan data yang diberikan.

Metode ini sering digunakan untuk analisis keranjang belanja, memungkinkan perusahaan untuk lebih memahami hubungan antara produk yang berbeda.

Transaksi 1	   
Transaksi 2	  
Transaksi 3	 
Transaksi 4	 
Transaksi 5	   
Transaksi 6	  
Transaksi 7	 
Transaksi 8	 

Jenis-jenis Unsupervised Learning

Apriori Algoritma

Apriori Algorithms dipopulerkan melalui analisis keranjang belanja yang mengarah ke berbagai mesin rekomendasi untuk platform musik dan toko online. Mereka biasa digunakan dalam kumpulan data transaksional untuk mengidentifikasi kumpulan barang yang sering, atau kumpulan barang, untuk mengidentifikasi kemungkinan mengonsumsi suatu produk mengingat konsumsi produk lain

Reinforcement Learning

Reinforcement learning (RL) bekerja melalui sebuah proses feedback, dan akan terus melakukan aktivitasnya sampai ia mencapai tujuannya. Jika ia mencapai tujuan maka ia akan mendapatkan reward. Proses ini akan terus berlangsung dengan tujuan besarnya adalah memaksimalkan reward yang didapat



Reinforcement Learning (lanjut)

Melalui gambar di atas bisa dilihat bahwa agen melakukan sebuah aksi, yang kemudian aksi itu akan diterapkan di lingkungannya (baik dunia luar atau berupa simulasi).

Kemudian akan dilihat apakah tujuannya tercapai atau tidak.

Jika tercapai maka ia akan mendapat reward, jika tidak maka ia mendapat hukuman (*punishment*) atau bisa juga disetting tidak terjadi apa-apa.

Setiap keputusan (aksi) yang ia ambil, maka ia berada di kondisi (*state*) yang baru. Begitu seterusnya sampai reward yang didapat maksimal.

Reinforcement Learning (lanjut)

- Dalam aplikasi nyata misalnya bermain mario bros, maka tujuan besarnya adalah mencapai skor tertinggi.
- Tentunya dengan tetap berusaha agar mario bros tidak pernah mati saat bermain.
- Atau jika kita mendesain robot yang bisa berjalan, maka tujuan besarnya adalah ia mencapai skor (reward) tertinggi juga. Setiap ia bisa berjalan normal (kriteria berjalan normal sudah kita tentukan), maka ia akan mendapat skor. Jika berjalannya sempurna, skornya sempurna, jika jatuh maka skornya minus. Perhitungan skor bersifat akumulatif. Begitu seterusnya.