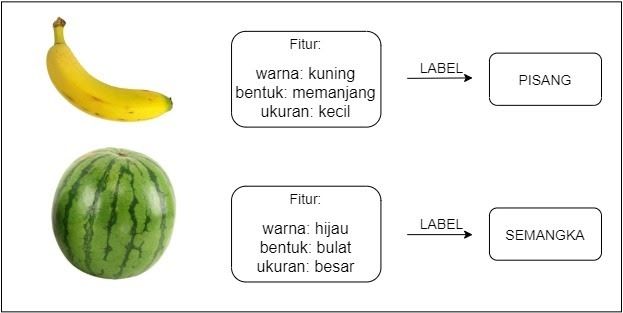
Pada supervised learning kita melatih sekumpulan data yang memiliki label. Label adalah pengenal dari sebuah data. Misal buah A memiliki atribut atau fitur berwarna hijau, bentuknya bulat, dan ukurannya sebesar bola sepak. Buah B atributnya berwarna kuning, bentuknya agak memanjang, dan besarnya segenggaman tangan. Buah yang memiliki atribut seperti yang disebutkan pada buah A dikenali sebagai semangka, sedangkan buah dengan atribut seperti disebutkan pada buah B dikenal sebagai pisang. Maka label dari buah A adalah semangka dan label buah B adalah pisang.



Sekumpulan data yang memiliki label biasanya digunakan untuk membuat model klasifikasi. Pada kasus di atas, Anda dapat menggunakan machine learning untuk membuat model klasifikasi buah. Model machine learning Anda akan mempelajari pola berdasarkan atribut dari data buah yang Anda miliki. Model juga belajar bahwa buah dengan atribut tertentu memiliki label tertentu, label A, misalnya. Dari pengalaman belajar tersebut, model machine learning Anda dapat digunakan untuk memprediksi data baru, dan menentukan apakah data baru ini termasuk buah dengan label A atau label B.

Algoritma supervised learning yang juga populer adalah regresi. Berikut adalah contoh permasalahan yang bisa diselesaikan dengan regresi. Anda memiliki data tentang penjualan properti di Kota Jakarta. Lantas Anda diminta untuk memprediksi harga properti di daerah Jakarta Selatan. Beberapa fitur penting yang berpengaruh terhadap harga properti biasanya adalah, lokasi, luas tanah, luas rumah, dan jumlah kamar. Dari data yang Anda miliki, model machine learning belajar bahwa properti dengan fitur-fitur berkelas, tentu akan memiliki harga tinggi, dan sebaliknya. Output dari model ini adalah sebuah fungsi yang dapat digunakan untuk memprediksi harga properti seperti yang diminta.

Pada masalah regresi kita memprediksi hasil dalam keluaran yang kontinyu berupa fungsi. Jika terdapat kontinuitas dalam keluaran model ML, maka dapat dipastikan itu adalah permasalahan regresi. Misalnya saat kita memprediksi harga rumah, model kita memprediksi harganya Rp. 499.999.000 atau Rp. 500.100.000, padahal harga seharusnya adalah Rp. 500.000.000. Perbedaan ini tak jadi masalah sebab hasil prediksi dalam regresi linear akan berada dalam range bilangan numerik tertentu yang mendekati nilai seharusnya. Inilah yang dimaksud dengan regresi memberikan keluaran yang kontinyu.

Sedangkan pada masalah klasifikasi, kita memprediksi hasil dalam keluaran diskrit. Misal, apakah email ini dapat diklasifikasikan sebagai spam atau bukan spam? Jelas bahwa tidak ada kontinuitas antara email spam dan bukan spam. Email spam tidak sama dengan bukan spam, serta tidak ada kategori antara email spam dan bukan spam.

Bagaimana, apakah Anda sudah bisa membedakan permasalahan klasifikasi dan regresi? Jangan khawatir, kita akan berlatih membuat model klasifikasi dan regresi sehingga Anda bisa lebih memahaminya.

Berikut adalah beberapa algoritma supervised learning yang penting Anda ketahui, dan akan dibahas di kelas Belajar Machine Learning untuk Pemula:

1. Linear Regression
2. Logistic Regression
3. Classification
4. Decision Trees
5. Support Vector Machines
6. Neural Networks.