Modul sebelumnya telah sedikit mengulas beberapa macam kategori machine learning yaitu, supervised, unsupervised, semi-supervised, dan reinforcement learning. Semi supervised merupakan gabungan dari supervised dan unsupervised. Sementara itu reinforcement learning dapat dianggap sebagai robot yang belajar untuk mendapatkan reward terbanyak.

Tiap kategori memiliki beberapa jenis algoritma. Algoritma yang termasuk supervised learning di antaranya adalah klasifikasi, regresi, decision tree, dan support vector machine. Sedangkan clustering, pendeteksian anomali (anomaly detection), dan pengurangan dimensi (dimension reduction) adalah jenis algoritma unsupervised learning.

Pada kelas Belajar Machine Learning untuk Pemula kita akan belajar tentang jenis-jenis algoritma dari kategori supervised dan unsupervised learning. Kenapa demikian? Sebabnya kini sebagian besar implementasi dari machine learning adalah berdasarkan algoritma supervised learning. Meskipun demikian, mayoritas data yang ada adalah data tanpa label.

Pada tahun 2016 seorang ilmuwan AI Yann LeCun memperkenalkan[“cake analogy”](https://syncedreview.com/2019/02/22/yann-lecun-cake-analogy-2-0/), sebuah analogi tentang machine learning yang dikenal luas hingga saat ini. Beliau menyatakan bahwa: “If intelligence is a cake, the bulk of the cake is unsupervised learning, the icing on the cake is supervised learning, and the cherry on the cake is reinforcement learning (RL).” Dalam Bahasa Indonesia, analogi ini artinya: “Jika sistem kecerdasan adalah sebuah kue, unsupervised learning akan menjadi kuenya, supervised learning adalah lapisan gula pada kue, dan reinforcement learning akan menjadi buah cherry pada kue”.



Dari analogi ini kita paham bahwa unsupervised memiliki bagian jauh lebih besar dibanding yang lainnya. Ilustrasi kue LeCun menyatakan bahwa data unsupervised berjumlah jutaan bits per sampel, yang mana itu sama dengan 100 hingga 100.000 kali lipat lebih besar dari data supervised. Mengagumkan, ya? Nah, masifnya ketersediaan data tanpa label ini menjadikan unsupervised learning memiliki potensi yang sangat besar untuk kita kembangkan.