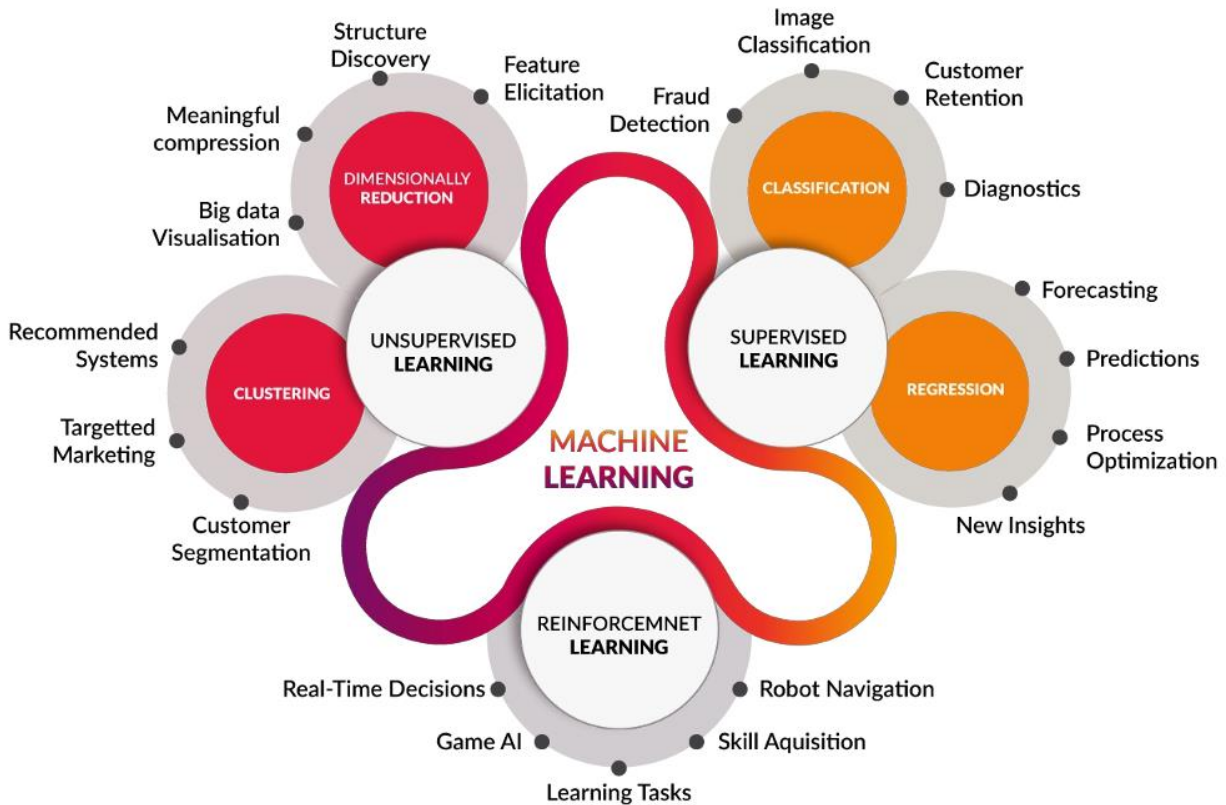


MACHINE LEARNING

Diciptakan oleh Samuel pada tahun 1959, istilah Machine Learning (ML) diberikan pada bidang studi yang memberikan kemampuan komputer untuk belajar tanpa diprogram secara eksplisit. Dengan kata lain, ML menyelidiki cara-cara di mana komputer dapat memperoleh pengetahuan secara langsung dari data untuk belajar memecahkan masalah.

Machine learning dan artificial intelligence adalah dua keyword yang paling banyak diperbincangkan saat ini. Namun, sebagian jurnalis, konsumen, dan pebisnis tidak terlalu mengerti kedua istilah tersebut. Dasarnya artificial intelligence adalah proses ‘melihat’, ‘mendengar’, dan ‘membaca’ data untuk mengambil keputusan. Jika dijabarkan secara lebih detail, artificial intelligence adalah proses replikasi kecerdasan manusia secara umum. Artificial intelligence digunakan bersamaan dengan machine learning. Sedangkan machine learning adalah salah satu topik pada ilmu komputer yang berbasis matematika dan statistika komputasi. Machine learning mengolah banyak data dan mempelajari pola dalam data untuk membuat prediksi di masa mendatang.

Algoritma machine learning digunakan untuk mengekstrak model yang didapat dari pengolahan data mentah (raw data) yang dapat digunakan untuk berbagai tugas dan tujuan. Cara kerja machine learning mirip dengan cara manusia belajar. Agar mesin, dalam konteks ini adalah komputer memiliki kemampuan belajar yang sama dengan manusia, maka perlu adanya proses klasifikasi sebelum menganalisis, menilai dan mengambil tindakan. Secara garis besar, algoritma machine learning dibagi menjadi tiga jenis, yaitu supervised learning, unsupervised learning, dan reinforcement learning. Sebelum menyelesaikan masalah menggunakan machine learning, kita harus memahami ketiga jenis machine learning tersebut karena algoritma-algoritma tersebut memiliki fungsi dan tujuan masing-masing. lalu apa yang dimaksud dengan supervised, unsupervised, dan reinforcement learning?



Gambar Machine Learning

(Sumber: <http://www.cognub.com/index.php/cognitive-platform/>)

Jenis-jenis Machine Learning antara lain:

1. Supervised

Supervised learning mengadopsi konsep pendekatan fungsi, dimana pada dasarnya algoritma dilatih agar dapat memilih fungsi-fungsi yang paling menggambarkan input dimana X tertentu membuat estimasi terbaik dari y . Namun, pada kenyataannya tidak sedikit orang yang kesulitan menemukan fungsi yang paling cocok. Hal ini karena sebenarnya algoritma bergantung pada asumsi yang digunakan. Jika ada asumsi yang tidak terpenuhi, tidak jarang hasil pengolahan data akan menimbulkan bias. Oleh karena itu, algoritma ini membutuhkan data latih yang benar sehingga sistem dapat mempelajari polanya dan regresi, klasifikasi, K-NN, Naive Bayes, Decision Trees, Regresi linier, Support Vector Machine, dan neural network.

Algoritma supervised learning adalah jenis machine learning yang paling umum. Dalam bahasa Indonesia, supervised learning diartikan sebagai pembelajaran diawasi. Istilah 'diawasi' ini muncul karena algoritma ini dirancang untuk belajar melalui contoh. Supervised learning memiliki beberapa kelebihan dan kekurangan. Keuntungan pertama, algoritma supervised learning adalah proses yang sederhana dan mudah dipahami. Selain itu algoritma ini juga powerful untuk klasifikasi. Data yang digunakan bukanlah data real time sehingga memerlukan data baru untuk memprediksi hasil. Kelemahan dari algoritma

ini adalah memerlukan waktu komputasi yang cukup panjang untuk pelatihan dan menggunakan algoritma yang lebih kompleks daripada algoritma unsupervised learning karena harus memberi label pada setiap input.

2. Unsupervised

Algoritma unsupervised learning adalah algoritma yang tidak membutuhkan data berlabel. Pada unsupervised learning, algoritma tidak membutuhkan data training. Algoritma ini digunakan dalam mendeteksi pola dan pemodelan deskriptif yang tidak membutuhkan kategori atau output berlabel yang menjadi dasar algoritma untuk mencari model yang tepat. Algoritma ini digunakan untuk clustering dan association rule. Keunggulan dari unsupervised learning adalah karena tidak membutuhkan label, algoritma lebih leluasa untuk mencari pola yang mungkin sebelumnya belum diketahui. Sedangkan kekurangan dari algoritma ini adalah sulitnya menemukan informasi dalam data karena tidak ada label dan lebih sulit untuk membandingkan output dengan inputnya.

Setelah memahami supervised dan unsupervised learning, ternyata tidak semua algoritma bisa dikategorikan sebagai supervised atau unsupervised learning. Algoritma tersebut bisa disebut dengan semi-supervised learning.

3. Reinforcement

Semi-supervised learning adalah algoritma yang tidak bisa dikategorikan ke dalam supervised atau unsupervised learning. Algoritma ini cocok digunakan untuk sejumlah data berukuran besar yang dibagi menjadi dua bagian yang diberi label dan tidak diberi label. Keuntungan dari semi-supervised learning adalah membutuhkan biaya yang lebih murah karena hanya setengah data yang diberi label dan tidak memerlukan tenaga ahli untuk mengolahnya.

Tujuan dari reinforcement learning adalah untuk menggunakan observasi yang dikumpulkan dari interaksi bersama lingkungan guna mengambil tindakan yang akan memaksimalkan output dan meminimalkan resiko. Algoritma ini akan terus belajar secara berulang-ulang. Dalam algoritma ini ada agen yang akan belajar dari interaksi dengan environment-nya. Untuk menghasilkan model, algoritma reinforcement learning melalui beberapa tahap antara lain agen mengamati data input, setelah itu agen melakukan suatu tindakan untuk mengambil keputusan. Setelah keputusan diambil, agen akan memperoleh 'reward' atau penguatan dari lingkungan. Lalu kembali mengamati input dan proses pengambilan keputusan kembali dilakukan namun dengan tambahan penguatan dari lingkungan sehingga hasil keputusan yang diambil lebih akurat.

Machine learning dan data science merupakan dua ilmu yang saling berkomplemen. Kedua ilmu ini saling melengkapi dan biasanya diaplikasikan secara bersamaan. Data science adalah ilmu yang sedang banyak dipelajari saat ini. Hal ini karena jumlah data yang semakin meningkat dan profesi data scientist sedang banyak dicari. Kedua ilmu ini dapat diterapkan di berbagai sektor, mulai dari finance, pendidikan, logistik, komunikasi, bahkan fashion. Kabar baiknya, ilmu data science dapat dipelajari oleh siapapun dengan latar belakang pendidikan apapun.

A. Machine Learning dan Statistik

Banyak ahli statistik percaya bahwa Automatic Interaction Detection (AID) awal dari pendekatan ML untuk menyelesaikan masalah statistik. AID dan keturunannya dianggap sebagai metode quasi-ML. Ini adalah teknik intensif komputer yang membutuhkan PC, syarat yang diperlukan untuk metode ML. Namun, bukanlah metode ML yang sebenarnya karena mereka menggunakan kriteria statistik secara eksplisit (misalnya, theta, Kuadrat dan uji-F) untuk pembelajaran. PC memungkinkan metode ML asli untuk belajar dengan meniru cara berpikir manusia. Jadi, digunakan istilah quasi. Mungkin istilah yang lebih tepat dan sugestif untuk prosedur tipe AID dan masalah statistik lainnya yang menggunakan PC adalah ML statistik.

Komunitas ML memiliki banyak kelompok yang bekerja pada data mining: neural networks, support vector machines, fuzzy logic, genetic algorithms and programming, information retrieval, knowledge acquisition, text processing, inductive logic programming, expert systems, and dynamic programming. Semuanya memiliki tujuan yang sama tetapi menyelesaikannya dengan alat dan teknik berbeda. Komunitas Statistik dan sub kelompok ML tidak memiliki pertukaran yang nyata untuk ide atau praktik terbaik. Diciptakan perbedaan tanpa perbedaan.

B. Machine Learning dan Matematika

Matematika yang menjadi dasar dari banyak metode pembelajaran mesin. Harapannya adalah bahwa seorang pembaca akan dapat mempelajari bentuk-bentuk dasar dari bahasa matematika, yang sekarang akan kita gunakan untuk mendeskripsikan dan mendiskusikan pembelajaran mesin.

Empat pilar machine learning:

1. Regression
2. Dimensionality reduction
3. Density estimation
4. Classification.

Mengilustrasikan bagaimana konsep matematika dapat digunakan untuk merancang algoritme pembelajaran mesin yang dapat digunakan untuk menyelesaikan tugas dalam lingkup empat pilar. Untuk memberikan serangkaian metode praktis yang memungkinkan pembaca untuk menerapkan pengetahuan yang diperoleh. Ini juga menyediakan gerbang ke literatur pembelajaran mesin yang lebih luas bagi pembaca yang sudah terbiasa dengan matematika.

Manusia berasumsi bahwa data dapat dibaca oleh komputer, dan direpresentasikan secara memadai dalam format numerik. Data diasumsikan berbentuk tabel di mana kita menganggap setiap baris tabel mewakili contoh atau contoh tertentu, dan setiap kolom menjadi fitur tertentu. Dalam beberapa tahun terakhir, pembelajaran mesin telah diterapkan pada banyak jenis data yang tidak jelas dalam format numerik tabular, misalnya urutan genom, konten teks dan gambar halaman web, dan grafik media sosial.

Banyak dari aspek matematika bergantung pada keahlian domain dan membutuhkan rekayasa yang cermat, dan aspek tersebut telah diletakkan di bawah ilmu data (Data Science).

[1] Bruce Ratner, *Statistical and Machine-Learning Data Mining: Techniques for Better Predictive Modeling and Analysis of Big Data*, Third Edition, CRC Press Taylor & Francis Group, 2017.

[2] Marc Peter Deisenroth, A. Aldo Faisal, Cheng Soon Ong, *Mathematics for Machine Learning*, Cambridge University Press, 2020.

[3] Samuel, A., Some studies in machine learning using the game of checkers, in Feigenbaum, E., and Feldman, J., Eds., *Computers and Thought*, McGraw-Hill, New York, 14–36, 1963.

[4]<https://www.dqlab.id/jenis-machine-learning-yang-perlu-diketahui#:~:text=Algoritma%20supervised%20learning%20adalah%20jenis,memiliki%20beberapa%20kelebihan%20dan%20kekurangan> Diakses Tanggal 30 April 2021