Supervised learning, juga dikenal sebagai pembelajaran mesin terawasi, adalah subkategori pembelajaran mesin dan kecerdasan buatan. Pembelajaran ini didefinisikan dengan penggunaan set data berlabel untuk melatih algoritma yang dapat mengklasifikasikan data atau memprediksi hasil secara akurat. Ketika data input dimasukkan ke dalam model, model akan menyesuaikan bobotnya hingga model tersebut sesuai, yang terjadi sebagai bagian dari proses validasi silang. Supervised learning membantu organisasi memecahkan berbagai masalah dunia nyata dalam skala besar, seperti mengklasifikasikan spam dalam folder terpisah dari inbox email

Konsep dasar dalam algoritma supervised learning adalah proses pembelajaran yang diawasi oleh guru, sedangkan dalam algoritma unsupervised learning, proses pembelajaran lebih bebas karena tidak ada pengawasan. Algoritma unsupervised learning lebih bebas dalam proses eksplorasi data karena tidak memiliki data label dan bisa mencari karakteristik data yang tersembunyi. Algoritma ini menggunakan titik data sebagai referensi untuk menemukan struktur dan pola yang ada di dalam data set.

Contoh supervised learning

Salah satu contoh pengaplikasian supervised learning yang paling menarik adalah memprediksi kondisi cuaca di lokasi tertentu. Untuk membuat prediksi cuaca yang benar, kita perlu memperhitungkan berbagai parameter, termasuk data suhu dari waktu ke waktu, curah hujan, angin, kelembaban, dan lain sebagainya. Metode yang tepat untuk memprediksi suhu adalah metode regresi dengan label output berupa data kontinu, sedangkan metode untuk memprediksi turunnya salju adalah metode klasifikasi binar.

Metode supervised learning yang paling populer adalah klasifikasi. Metode ini digunakan e-commerce untuk memprediksi sentimen teks dari tweets atau ulasan produk mereka. Jika tweets atau ulasan produk masuk ke dalam kelas positif, maka dapat diartikan bahwa pelanggan puas dengan produk tersebut.

Contoh unsupervised learning

Contoh pengaplikasian algoritma unsupervised learning adalah segmentasi pelanggan, mengurangi kompleksitas suatu masalah, dan memilih fitur yang tepat. Segmentasi pelanggan membutuhkan metode clustering. Metode ini termasuk algoritma unsupervised learning yang bertujuan untuk menemukan kelompok atau cluster alami di dalam data input. Salah satu pendekatan umum dalam clustering adalah membagi titik data sedemikian rupa sehingga setiap titik data yang memiliki kemiripan masuk ke dalam grup yang sama.

Metode clustering biasanya digunakan untuk menentukan segmen pelanggan dalam data pemasaran. Jika tim pemasaran memiliki data segmentasi pelanggan, maka tim pemasaran dapat melakukan pendekatan yang tepat ke setiap segmen pelanggan. Pengurangan dimensi adalah salah satu teknik algoritma unsupervised learning yang umum digunakan dengan tujuan untuk mengurangi jumlah variabel acak yang sedang dipertimbangkan.

Video:  
<https://youtu.be/W01tIRP_Rqs>