

# Scikit-Learn

Scikit-Learn adalah perpustakaan Python populer untuk belajar mesin (*machine learning*), dibangun di atas SciPy (modul sains dan komputasi ilmiah). Perpustakaan ini gratis, open-source, dan berlisensi BSD 3-Clause, sehingga bisa digunakan siapa saja, termasuk untuk keperluan komersial.

## Supervised Learning di Scikit-Learn

Ini adalah teknik di mana model belajar dari data yang sudah memiliki "jawaban benar". Contohnya:

Contoh Penerapan di Scikit-Learn:

- **Klasifikasi:**

Misalnya, memprediksi apakah sebuah email adalah *spam* atau bukan. Contoh algoritma:

- **Decision Tree:** Seperti pohon keputusan untuk klasifikasi, misalnya menentukan apakah pelanggan akan membeli produk berdasarkan riwayat belanja.
- **SVM (Support Vector Machine):** Cocok untuk tugas kompleks seperti membedakan gambar kucing dan anjing.

- **Regresi:**

Misalnya, memprediksi harga rumah berdasarkan ukuran tanah, lokasi, atau jumlah kamar.

Contoh algoritma:

- **Linear Regression:** Sederhana dan efektif untuk memprediksi nilai numerik (seperti harga).

## Unsupervised Learning di Scikit-Learn

Di sini, model belajar mencari pola dari data yang *tidak memiliki label*. Contohnya:

- **Clustering (Pengelompokan):**

Misalnya, mengelompokkan pelanggan berdasarkan kebiasaan belanja. Contoh algoritma:

- **K-Means:** Membagi data ke dalam kelompok (misalnya 3 kelompok pelanggan: hemat, menengah, tinggi).
- **DBSCAN:** Cocok untuk data dengan kepadatan tidak merata, seperti mendeteksi area ramai pengunjung di peta.

- **Reduksi Dimensi:**

Teknik untuk menyederhanakan data tanpa kehilangan informasi penting. Contoh algoritma:

- **PCA (Principal Component Analysis):** Mengurangi fitur data (misalnya dari 100 fitur jadi 2 fitur) untuk visualisasi atau mempercepat pemrosesan.