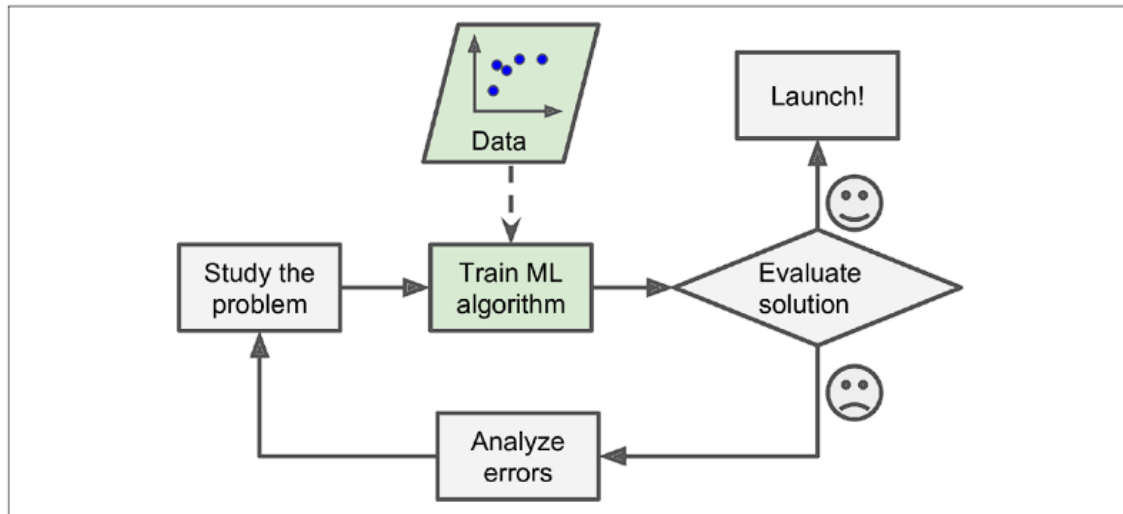


The Machine Learning Landscape

Bab 1 dari buku *Hands-On Machine Learning with Scikit-Learn and TensorFlow* ini memberikan gambaran menyeluruh tentang konsep dasar *machine learning* (ML) dan menggarisbawahi pentingnya teknologi ini dalam berbagai bidang.



A. Definisi dan Pendekatan ML :

- *Machine learning* didefinisikan sebagai suatu metode di mana sebuah sistem komputer dapat belajar dari data tanpa perlu diprogram secara eksplisit.
- Pendekatan ML muncul karena banyak masalah di dunia nyata terlalu rumit untuk diselesaikan hanya dengan aturan statis yang dibuat oleh manusia; oleh karena itu, mesin perlu menemukan polanya sendiri dari data yang tersedia.

B. Kategori Utama *Machine Learning* :

ML diklasifikasikan ke dalam beberapa kategori berdasarkan cara belajar dan tipe data yang digunakan:

- Supervised Learning: Melibatkan pelatihan dengan data berlabel.
- Unsupervised Learning: Digunakan untuk menemukan struktur atau pola tersembunyi dalam data tanpa label.
- Semi-supervised Learning dan Reinforcement Learning juga termasuk dalam klasifikasi umum.

C. Metode Pelatihan dan Pemodelan

Terdapat perbedaan dalam cara model dilatih dan dibangun:

- Batch Learning: Seluruh pelatihan dilakukan sekaligus menggunakan seluruh *dataset*.
- Online Learning: Model diperbarui secara bertahap dari aliran data yang terus-menerus.
- Instance-based Learning: Model membuat keputusan dengan mengandalkan penyimpanan data (misalnya, K-Nearest Neighbors/KNN).

- Model-based Learning: Model bekerja dengan membangun model matematis dari data.

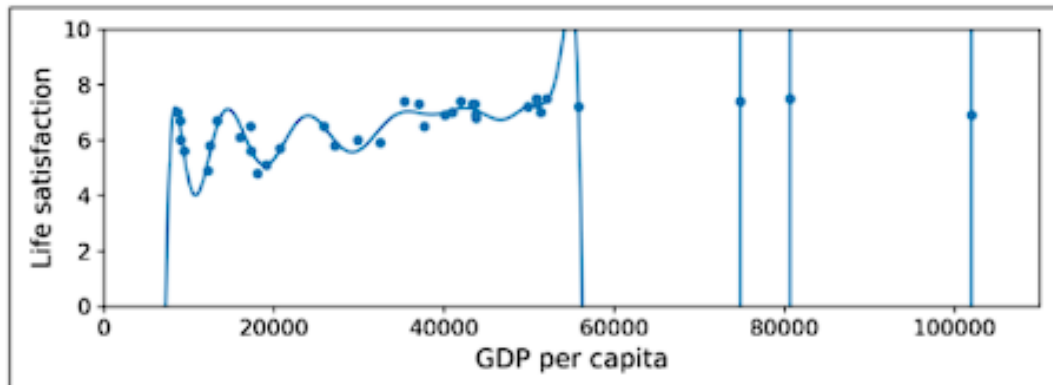


Figure 1-22. Overfitting the training data

D. Tantangan dan Proses

Bab ini juga menguraikan tantangan utama yang dihadapi dalam proyek ML, seperti kualitas data yang buruk, bias, data yang tidak lengkap, serta isu *underfitting* (model terlalu sederhana) dan *overfitting* (model terlalu kompleks).