Supervised learning

Training dataset contains pairs of input and corresponding output values
Goal: create an accurate predictor for future data

Classification

Output variable is discrete: label or binary

Regression

Output variable is real-valued: e.g., value of a physical property

Unsupervised learning

Training dataset contains input with no corresponding labels or output
Goal: identify undetected patterns in the data

Clustering

Group data points according to their similarity

Principal component

Reduce dimensionality of data and identify trends