

Nama : Ferza Reyaldi  
NIM : 09021281924060  
Mata Kuliah : Pengenalan Pola

## RESUME KULIAH UMUM

### *International Public Lecture Pattern Recognition*

#### *What is Machine Learning?*

Machine learning adalah cabang ilmu kecerdasan buatan yang membuat komputer mampu 'belajar' tanpa diprogram secara eksplisit.

Menurut Tom Mitchell (1998), Machine learning adalah ilmu algoritma yang meningkatkan *performance* (P) pada suatu *task* (T) dengan menambah *experience* (E). Contoh:

- T: mengidentifikasi seseorang melalui gambar wajah
- P: persentase kebenaran seseorang diidentifikasi
- E: database gambar wajah

#### *Learning System Model*

- *Statistical Approach*: mengambil kesimpulan dari sampel, mencirikan hubungan antara data dan variabel hasil.
- *Computer Science Approach*: mengembangkan model untuk menyelesaikan *optimization problem*. Mengevaluasi model prediksi.

#### *Optimization Problem*

Masalah yang hanya fokus mencari 2 nilai, nilai maksimum dan minimum. Yaitu:

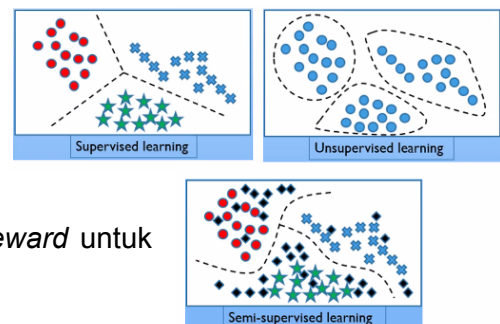
- combinatorial optimization: greedy search (dijkstra, kruskal)
- Convex optimization: gradient descent (neural network, deep learning)
- Constrained optimization: linear programming

#### *When to use Machine Learning?*

- Seorang pakar tidak cukup
- Layanan yang tidak bisa dijelaskan dengan kemampuan manusia
- Layanan personalisasi (sistem rekomendasi)
- Jumlah data yang sangat besar
- Meningkatnya kemampuan komputasi
- Algoritma terobosan baru dari para peneliti

#### *Types of learning*

- Supervised learning: training data memiliki label
- Unsupervised learning: training data tidak memiliki label
- Semi-supervised learning: sebagian training data tidak memiliki label
- Reinforcement learning: menggunakan sistem *reward* untuk serangkaian aksi mesin.



### How to measure performance?

- Supervised: ditinjau dari nilai error (prediksi yang salah) dari testing set.

$$error = \frac{1}{N} \sum_{n=1}^N [y_n \neq g(x_n)]$$

- Unsupervised: ditinjau dari minimum quantization error, minimum distance, maximum likelihood estimation.

### Data Pre-processing

Data pre-processing adalah proses yang dilewati sebelum melakukan pembuatan model dalam machine learning. Pada tahap ini dilakukan cleaning data, normalizing data, feature selection, feature engineering, memvalidasi data dan lain-lain.

### Algorithm

- *KNN Classification*: algoritma yang bertujuan mengklasifikasikan data, data yang dimiliki label berupa kategorikal. Data input dikelompokkan dengan mengecek jaraknya dengan  $k$  instance terdekat. Misal  $k = 3$ , jika data input dekat dengan 2 data label merah dan 1 data label biru, maka data inputan dapat kita klasifikasikan sebagai data merah. Nilai  $k$  disarankan menggunakan bilangan ganjil lebih dari 1 untuk mencegah terjadi ambiguitas dalam penentuan label data.
- *KNN Regression*: algoritma supervised yang memprediksi nilai output dari suatu input, data training memiliki label berupa numerik. Algoritma ini memberikan pengenalan ke  $K$ -terdekat regresi tetangga dalam memprediksi nilai output.
- *K-Means Clustering*: data dikelompokkan ke dalam  $k$  centroid yang diletakkan secara random oleh mesin. Data akan dikelompokkan ke dalam centroid terdekat.

### Advanced Machine Learning

Beberapa bidang lanjutan dari machine learning antara lain:

- *Computer vision*: image segmentation, face recognition, object detection, object tracking, autonomous car, dan lain-lain.
- *Natural language processing*: world translation, sentiment analysis, speech recognition, digital assistant, dan lain-lain.
- *Internet of Things (IoT)*: traffic density estimation, fall detection, heart arrhythmia detection, smart factory, smart car, dan lain-lain.