

Praktikum Machine Learning

Maret 2022

1 Supervised Learning

Berikut adalah beberapa algoritma *supervised learning* yang akan digunakan

1.1 Naive Bayes

Naive Bayes Classifier adalah algoritma klasifikasi yang bergantung pada probabilitas. Cek file *notebook* `Naive Bayes using sklearn.ipynb`. Lakukan perintah untuk mengklasifikasikan data seperti pada contoh, namun dengan data `predictive_maintenance.csv`. Lakukan evaluasi dengan *confusion matrix* dan *plot* seperti pada contoh.

1.2 Decision Tree

Decision Tree Classifier adalah algoritma klasifikasi yang bergantung pada percabangan dari setiap kondisi *features*. Cek file *notebook* `Decision Trees`. Lakukan perintah untuk mengklasifikasikan data seperti pada contoh, namun dengan data `predictive_maintenance.csv`. *Features* tidak perlu distandardisasi atau dinormalisasi pada algoritma ini, karena pengelompokkan tidak berdasarkan jarak pada ruang *features*. Lakukan evaluasi dengan *confusion matrix* dan *plot* seperti pada contoh.

1.3 K-nearest Neighbours

K nearest neighbour adalah metode dimana sebuah data akan di-*match* dengan hasil *training* pada ruang *features*. Cek file *notebook* `KNN.ipynb`. Lakukan perintah untuk mengklasifikasikan data seperti pada contoh, namun dengan data `predictive_maintenance.csv`. Lakukan evaluasi dengan *confusion matrix* dan *plot* seperti pada contoh.

1.4 Artificial Neural Network

Artificial Neural Network (ANN) adalah metode *machine learning* (atau lebih spesifiknya *deep learning*) yang menirukan sistem syaraf di otak. Sistem ini belajar dengan mengenali sampel train. Cek file *notebook* `ANN.ipynb`. Dalam neural network, dikenal istilah epoch, yaitu satu siklus training dataset. Dikenal juga istilah neuron dan layer, dimana berfungsi untuk menghitung *weight* dan mengeluarkan output dengan fungsi aktivasi.

Lakukan perintah untuk mengklasifikasikan data seperti pada contoh, namun dengan data `predictive_maintenance.csv`. Lakukan evaluasi dengan *confusion matrix* dan akurasi seperti pada contoh. Variasikan nilai *epoch* menjadi 10,20,30,40,dan 50 epoch. Apakah perbedaannya?

1.5 Komparasi

Dari semua algoritma diatas, manakah yang akan anda pilih sebagai preferensi untuk melakukan *fitting* machine learning? Sebutkan pertimbangannya. Jelaskan pada cell markdown saja.