

IF4071Pembelajaran Mesin

Tugas Kecil II: Eksplorasi scikit-learn pada Jupyter Notebook

23Oktober 2017

Jupyter Notebook (<http://jupyter.org/>) memudahkan kita untuk membuat dan men-share dokumen yang merupakan gabungan dari live code, equation, visualisasi dan catatan. Jupyter dapat digunakan untuk visualisasi, pembersihan dan data transformasi, statistical model dan machine learning. Scikit-learn merupakan library machine learning pada bahasa python.

1. Lakukan eksplorasi scikit learn pada Jupiter Netbookdan bacalah dokumentasinya:<https://jupyter-notebook.readthedocs.io/en/stable/notebook.html>
<http://scikit-learn.org/stable/documentation.html>
2. Tulislah script dalam bahasa python pada satu notebook untuk melakukan task berikut ini:
 - a. Membaca dataset standar iris dan dataset play-tennis (dataset eksternal dalam format csv). Gunakanlah sklearn.datasets untuk membaca dataset standar. Untuk membaca dataset csv, gunakanlah Python Data Analysis Library <http://pandas.pydata.org/>
 - b. Melakukan pembelajaran: DecisionTree (<http://scikit-learn.org/stable/modules/tree.html>), dan Artificial Neural Network (http://scikit-learn.org/stable/modules/neural_networks_supervised.html) untuk dataset iris dengan skema full-training, dan menampilkan modelnya.
 - c. Melakukan pembelajaran DecisionTree, dan ANN untuk dataset iris dengan skema split train 90% dan test 10%, dan menampilkan kinerja serta confusion matrixnya.
 - d. Melakukan pembelajaran DecisionTree, dan ANN untuk dataset iris dengan skema 10-fold cross validation, dan menampilkan kinerjanya.
 - e. Menyimpan (save) model/hipotesis hasil pembelajaran ke sebuah file eksternal
 - f. Membaca (read)model/hipotesis dari file eksternal
 - g. Membuat instance baru dengan memberi nilai untuk setiap atribut
 - h. Melakukan klasifikasi dengan memanfaatkan model/hipotesis DecisionTree, dan ANN pada instance pada g.
 - i. Melakukan analisis fitur dengan visualisasi hubungan setiap fitur dengan kelas. Gunakanlah seaborn (<https://seaborn.pydata.org/>).
4. Tugas dikerjakan berkelompok, dan 1 kelompok terdiri atas 3 mahasiswa (boleh gabungan mahasiswa K1 dan K2). Kelompok yang sama akan digunakan untuk eksperimen.
5. Tugas dikumpulkan berupa file .ipynb dan pdf dengan isi laporan sesuai ketentuan butir tiga. Hanya salah satu anggota kelompok saja yang *uplad* file tugas pada website ini.
6. Penamaan file yang dikumpulkan: Tucil2_[NIM salah satu anggota].zip (misal: Tucil2_13515001.zip yang berisi Tucil2_13515001.ipynb dan Tucil2_13515001.pdf).
7. Pengumpulan yang terlambat tidak diperbolehkan, batas akhir adalah hari Minggu, 29Oktober 2017 jam 23.59 (waktu situs ini).