

TUGAS EKSPLORASI SCIKIT LEARN PADA JUPYTER NOTEBOOK INTELIGENSI BUATAN T.A. 2018/2019

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 Netbook dan bacalah dokumentasinya :

<https://jupyter-notebook.readthedocs.io/en/stable/notebook.html>

<http://scikit-learn.org/stable/documentation.html>

2. Proses Instalasi

- a. Instalasi di windows

- Cara paling mudah adalah menggunakan **Anaconda** yang dapat diunduh pada laman <https://www.anaconda.com/download/#download>. Setelah proses instalasi selesai dilakukan, maka carilah “Anaconda Prompt” selanjutnya ketikkan command line:

```
jupyter notebook
```

- Jupyter akan otomatis muncul di browser
- Untuk instalasi berbagai library yang diperlukan, buka Anaconda Prompt kembali dari awal, dan gunakan command line berikut:

```
conda install pandas
```

```
conda install scikit-learn
```

- b. Instalasi di Linux

- Untuk melakukan instalasi anaconda pada sistem operasi linux dapat melalui terminal dan menggunakan command line:

```
pip3 install jupyter
```

- Sedangkan untuk menjalankannya bisa menggunakan command line:

```
jupyter notebook
```

- Untuk instalasi package library yang dibutuhkan bisa menggunakan fasilitas pip3, dengan menggunakan command line:

```
pip3 install pandas
```

```
pip3 install scikit-learn
```

3. 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:
NaiveBayes (http://scikit-learn.org/stable/modules/naive_bayes.html),
DecisionTree ID3 (<http://scikit-learn.org/stable/modules/tree.html>),
kNN (<http://scikit-learn.org/stable/modules/neighbors.html>), dan
Neural Network MLP (http://scikit-learn.org/stable/modules/neural_networks_supervised.html)
untuk dataset iris dengan skema full-training, dan menampilkan modelnya.
 - c. Melakukan pembelajaran NaïveBayes, DecisionTree, kNN, dan MLP untuk dataset iris dengan skema split train 90% dan test 10%, dan menampilkan kinerja serta confusion matrixnya.
 - d. Melakukan pembelajaran NaïveBayes, DecisionTree, kNN, dan MLP 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/hipotesisNaïveBayes, DecisionTree, dan kNN dan instance pada g.
- 4. Tugas dikerjakan berkelompok, dan 1 kelompok terdiri atas 2 mahasiswa.
 - 5. Tugas dikumpulkan berupa hasil download notebook dalam dua format yaitu file .ipynb dan pdf. Hanya salah satu anggota kelompok saja yang melakukan *upload* file tugas pada kuliah.itera.ac.id. Penamaan file yang dikumpulkan: Tubes_[NIM1_NIM2].zip (misal: Tubes_13515001_13515002.zip yang berisi file dengan ekstensi .ipynb dan .pdf).
 - 6. Pengumpulan yang terlambat tidak diperbolehkan, batas akhir adalah hari **Jumat, 14 Desember 2018 pukul 23.55 WIB** (waktu sistem).