



Minggu ke-4

Praktikum Validation Model of Classification

Entin Martiana, Ali Ridho Barakbah

Knowledge Engineering Research Group
Department of Information and Computer Engineering
Politeknik Elektronika Negeri Surabaya



Melakukan Normalisasi Data dengan metode Min-Max dengan acuan Data Training

```
train_data = np.array(xtrain)[: ,1:-1]
newmin=0
newmax=1
mindata = train_data.min()
maxdata = train_data.max()
train_data = ((train_data-mindata)*(newmax-newmin)/(maxdata-mindata))+newmin
test_data = np.array(xtest)[: ,1:-1]
test_data = ((test_data-mindata)*(newmax-newmin)/(maxdata-mindata))+newmin
```



Cara Memanggil Library untuk Metode Hold-out

```
from sklearn.model_selection import train_test_split

dataset = pd.read_csv('ruspini.csv')
datalabel = dataset.loc[:,['CLASS']]
xtrain, xtest, ytrain, ytest = train_test_split(dataset, datalabel, test_size = 0.30, random_state=100)
```



Cara Memanggil Library untuk Metode k-Fold

```
from sklearn.model_selection import KFold

dataset = pd.read_csv('ruspini.csv')
kf=KFold(n_splits=10,random_state=0,shuffle=True)
p=0
for train_index, test_index in kf.split(dataset):
    p=p+1
    xtrain=dataset.loc[train_index]
    xtest=dataset.loc[test_index]
    ytrain=xtrain.loc[:,['CLASS']]
    ytest=xtest.loc[:,['CLASS']]
```



Cara Memanggil Library untuk Metode LOO

```
from sklearn.model_selection import LeaveOneOut

dataset = pd.read_csv('ruspini.csv')
loo=LeaveOneOut()
n=loo.get_n_splits(dataset)
for xtrain, xtest in loo.split(dataset):
    xtrain = dataset.filter(items=xtrain, axis=0)
    xtest = dataset.filter(items=xtest, axis=0)
ytrain=xtrain.loc[:,['CLASS']]
ytest=xtest.loc[:,['CLASS']]
```



Menghitung Error Ratio

```
precision_ratio=kNN.score(test_data,test_label)
error_ratio = 1- precision_ratio
print(error_ratio)
```



Assignment #3 – Validation Model

1. dataset \leftarrow milk.csv
2. Lakukan validation Model dengan metode:
 - a. Hold-out Method (70%-30%)
 - b. K-Fold (k=10)
 - c. LOO
3. train_data \leftarrow lakukan normalisasi pada train_data dengan Min-Max 0-1 (catat nilai min dan max setiap atribut)
4. test_data \leftarrow lakukan normalisasi pada train_data dengan min-max pada Langkah 5
5. Lakukan klasifikasi k-NN (k=3) untuk masing-masing pendekatan validasi dan hitunglah error ratio-nya



Pengumpulan Tugas

- Buatlah coding dengan Bahasa pemrograman/tools apapun untuk semua assignment
- Buatlah laporan dalam slide ppt. Laporan terdiri dari screenshot coding dan hasil running untuk setiap assignment.
- Simpan laporan dalam file pdf dengan format penamaan:
MLSDT_M4_NRP_namadepan.pdf
- Upload file tersebut ke ETHOL
- Deadline upload: Kamis, 5 Oktober 2023

