



Minggu ke-6

Praktikum Decission Tree

Entin Martiana, Ali Ridho Barakbah, Renovita

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



Klasifikasi dengan Decision Tree

```
from sklearn.tree import DecisionTreeClassifier  
  
...  
  
dtc=DecisionTreeClassifier()  
dtc.fit(train_data, train_label)  
class_result=dtc.predict(test_data)  
  
...  
  
acc=dtc.score(train_data, train_label)
```



Menampilkan Gambar Decision Tree

```
from sklearn import tree
import graphviz

...

dot_data = tree.export_graphviz(dtc, out_file=None,
                               feature_names=train_data.columns.values)

graph = graphviz.Source(dot_data, format="png")
graph.render(view=True)
```



Contoh Klasifikasi pada Ruspini Dataset

```
import pandas as pd
from sklearn import tree
from sklearn.tree import DecisionTreeClassifier
import graphviz

dataset = pd.read_csv('ruspini.csv')

train_data=dataset[['X', 'Y']]
train_label=dataset[['CLASS']]

dtc=DecisionTreeClassifier()
dtc.fit(train_data, train_label)
test_data=[[130, 45]]
class_result=dtc.predict(test_data)
print('Class = ', class_result)

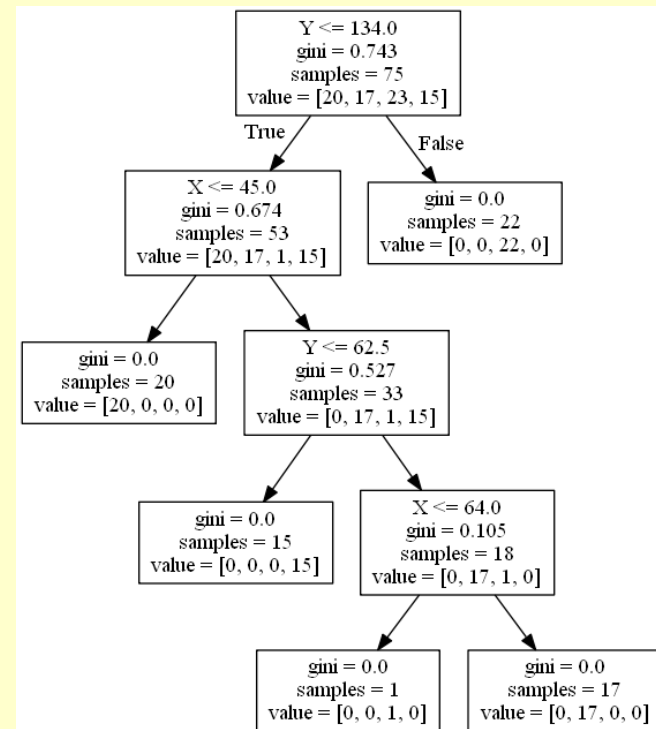
acc=dtc.score(train_data, train_label)
err=round((1-acc)*100, 2)
print('\n\nError ratio = ', err, '%')

dot_data = tree.export_graphviz(dtc, out_file=None,
feature_names=train_data.columns.values)

graph = graphviz.Source(dot_data, format="png")
graph.render( view=True)
```

Class = [4]

Error ratio = 0.0 %



Assignment #4 – Validation Model

1. `train_dataset` \leftarrow `milk_training.csv`
2. `train_data` \leftarrow ambil `train_dataset` kolom fitur (pH, Temperature, Taste, Odor, Fat, Turbidity, Colour).
3. `test_dataset` \leftarrow `milk_testing.csv`
4. `test_data` \leftarrow ambil `test_dataset` kolom fitur (pH, Temperature, Taste, Odor, Fat, Turbidity, Colour).
5. `train_label` \leftarrow ambil `train_data` kolom kelas (Grade)
6. `test_label` \leftarrow ambil `test_data` kolom kelas (Grade)
7. Lakukan klasifikasi `test_data` terhadap `train_data` dengan Decision Tree, dan berapakah error rasionya?
8. Tampilkan hirarki dari Decision Tree
9. Bandingkan dengan hasil dari k-NN dan Bayesian



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:
MLLJ_M6_NRP_namadepan.pdf
- Upload file tersebut ke ETHOL
- Deadline upload: Minggu, 12 Oktober 2023

