

Application 4

Supervised Machine Learning

Iris Dataset with Decision Tree Plotting

- In this application we remove one entry from each label of iris dataset and train with the remaining entries.
- And we apply predictions based on Decision tree with that removed entries.
- We plot the decision tree using pydot library.

Consider below characteristics of Machine Learning Application :

Classifier :	Decision Tree
DataSet :	Iris Dataset
Features :	Sepal Width, Sepal Length, Petal Width, Petal Length
Labels :	Versicolor, Setosa , Virginica
Training Dataset :	147 Entries
Testing Dataset :	3 Entries

```
1 from sklearn.datasets import load_iris
2 import numpy as np
3 from sklearn import tree
4
5 iris = load_iris()
6
7 print("Feature names of iris data set")
8 print(iris.feature_names)
9
10 print("Target names of iris data set")
11 print(iris.target_names)
12
13 # Indices of removed elements
14 test_index = [1,51,101]
15
16 # Training data with removed elements
17 train_target = np.delete(iris.target,test_index)
18 train_data = np.delete(iris.data,test_index,axis=0)
19
20 # Testing data for testing on training data
21 test_target = iris.target[test_index]
22 test_data = iris.data[test_index]
23
```

```

24 # form decision tree classifier
25 classifier = tree.DecisionTreeClassifier()
26
27 # Apply training data to form tree
28 classifier.fit(train_data,train_target)
29
30 print("Values that we removed for testing")
31 print(test_target)
32
33 print("Result of testing")
34 print(classifier.predict(test_data))
35
36 #visualization
37 from sklearn.externals.six import StringIO
38 import pydot
39
40 dot_data = StringIO()
41 tree.export_graphviz(classifier,out_file=dot_data,feature_names=iris.feature_names,
42                     class_names=iris.target_names,filled=True,rounded=True,impurity=False)
43
44 graph = pydot.graph_from_dot_data(dot_data.getvalue())
45 graph[0].write_pdf("Marvellous.pdf")
46

```

When we execute the above application which create **Marvellous.pdf** file in current directory.

That PDF file contains visual representation of our Decision tree.

