作业:使用决策树预测隐形眼镜类型

1、数据集：<http://archive.ics.uci.edu/ml/machine-learning-databases/lenses/>

2、利用scikit-learn中的DecisionTreeClassfier进行分类,输出准确率,决策树图形

scikit-learn中决策树的可视化一般需要安装graphviz。主要包括graphviz的安装和python的graphviz插件的安装。

第一步是安装graphviz。下载地址在：http://www.graphviz.org/。如果你是linux，可以用apt-get或者yum的方法安装。如果是windows，就在官网下载msi文件安装。无论是linux还是windows，装完后都要设置环境变量，将graphviz的bin目录加到PATH，比如我是windows，将C:/Program Files (x86)/Graphviz2.38/bin/加入了PATH

第二步是安装python插件graphviz： pip install graphviz

第三步是安装python插件pydotplus。这个没有什么好说的: pip install pydotplus

from sklearn import tree

import pydotplus

from IPython.display import Image

import sys

import os

os.environ["PATH"] += os.pathsep + 'D:/Program Files/Graphviz2.38/bin/'

from sklearn.datasets import load\_iris

iris = load\_iris()

clf = tree.DecisionTreeClassifier(criterion='entropy')

clf = clf.fit(iris.data, iris.target)

clf.predict(iris.data[:1, :])

dot\_data = tree.export\_graphviz(clf, out\_file=None)

graph = pydotplus.graph\_from\_dot\_data(dot\_data)

graph.write\_pdf("c:\iris.pdf")

# 训练模型，限制树的最大深度4

clf = tree.DecisionTreeClassifier(max\_depth=3)

#拟合模型

clf.fit(X, y)

dot\_data = tree.export\_graphviz(clf, out\_file=None)

graph = pydotplus.graph\_from\_dot\_data(dot\_data)

graph.write\_pdf("c:\iris1.pdf")