《人工智能与物理学》

第3次作业

作业规范：

同时提交如下两个电子文件：
（1）一个PDF或Word文档描述你的想法和结果（可以选择把对应代码
　　　复制在这里，但是（2）中规定的可运行的代码仍然必不可少）
（2）附上可单独运行的代码

注1: 作业提交**截止时间**：11月6日周日晚24:00之前，过时不候。
注2: **不准使用**神经网络或相关深度学习模型。其余任何你喜欢的算法都可以。
注3: 如果你想压缩文件，请用.zip，不要用.rar。

Homework 3-1

A typical regression task.

- Each item has eleven features (named 'Feature1', 'Feature2', …, 'Feature11' respectively) and one integer number 'Yvalue' indicating the quality of an item.

- The dataset has been split into two parts: Homework-3-1-train_data.csv and Homework-3-1-test_data.csv.

◆Please develop a regression model and use the train_data to train your model, and then apply your model to predict the quality ('Yvalue') of items in test_data.

◆RMSE (root mean squared error) <u>must</u> be used to assess the inaccuracy of your predictions (see the example code below, don't forget **0.5). **Your goal is to make RMSE as small as possible.**

```python
from sklearn.metrics import mean_squared_error


test_rmse = mean_squared_error(test_pred,y_test)**0.5
print("Prediction RMSE: {}%".format(float(test_rmse)))
```

Homework 3-2

A typical classification task.

- Each item has five features (named 'Feature1', 'Feature2', ..., 'Feature5' respectively) and one Boolean number (i.e., 0 or 1) indicating the 'Class' of an item.

- The dataset has been split into two parts: Homework-3-2-train_data.csv and Homework-3-2-test_data.csv.

◆Please develop a classifier and use the train_data to train your model, and then apply your model to predict the 'Class' of items in test_data.

◆AUC (area under ROC curve) <u>must</u> be used to assess the performance of your model (see the next page for **tips**)

<u>Your goal is to make AUC as large as possible</u>.

Homework 3-2

Tips:

• There is a function defined for drawing AUC curve (see AUC_code.py). You can copy it into your code, and easily get the following figure.