

Artificial Intelligence

Homework#1

目標：

- 利用Iris dataset測試K-NN分類器，並求出其分類率。

資料描述：

- 鳶尾花資料為機器學習領域中，常被用來驗算演算法優劣的資料庫。數據庫中包含三種不同鳶尾花(山鳶尾、變色鳶尾以及維吉尼亞鳶尾)。每種花有50筆樣本，每筆樣本以花萼長度、花萼寬度、花瓣長度以及花瓣寬度四種數值作為特徵，進行後續定量分析。
- MATLAB讀取鳶尾花資料後會產生150×5的陣列，其中第5行(5th column)為資料的類別標籤。

作業內容：

1. 利用兩兩特徵畫出散佈圖(Scatter plot)，共6張圖。
2. 利用Iris dataset測試K-NN分類器，並列出所有可能之特徵組合(共15種組合)的分類率。
3. 討論分類率與Scatter plot之關係，並附上心得。

Note:

- 第一次先將各類別資料中的前半資料當作測試資料(Training data)，剩下的後半資料當作測試資料(Testing data)，求得一個分類率；之後再將Training data和Testing data互換，求得第二個分類率，再將兩分類率平均。
- K-NN中之K值分別取1和3。

Deadline：2020/3/31(two weeks)

聯絡方式：

助教：蕭玉聰

Email：chariestas520@gmail.com

研究室：綜科 707-3