類神經網路簡介HW4

Q36111281 張偉治

1. 原始資料處理

1.1 分類資料轉換

將七類的豆子名稱分別對應到數字1到7

1.2 移除離群值

使用四分位數的方式移除離群值，其中參數n=3，透過該方法總共移除12%的資料。

一張含有 文字 的圖片

自動產生的描述

1.3 填補空值

一張含有 桌 的圖片

自動產生的描述

2. SVM調整參數

一張含有 文字 的圖片

自動產生的描述

分別使用兩種SVM模型：linear和rbf，其中linear的部分又使用兩種不的C值，區分不同模型的複雜度。

最後對應的準確率分別為0.903、0.902、0.300可以發現linear模型的SVM對該資料集具有較好的精確度。

3. 其他分類器

這次除使用NN外，另外還使用bagging的技巧，最後分別使用較寬與較深的NN模型對bagging後的資料進行預測，發現較寬的NN2對bagging具有相對較好的準確度。