

2019 ECT Project

請用之前作業處理好 fifa19 的 data_new.csv 進行資料分析, 過程中對所有重要程式步驟進行截圖並加以說明, 越詳盡越好。

- 總共有 5 大題與 1 題加分題(滿分 120 分)

第 1- 4 題

- 從 data_new.csv 選取 Crossing ~ GKReflexes 欄位(共 34 個屬性)
- 計算以上所有欄位的平均
- 加標籤(大於平均: 'Above-average Players', 小於平均: 'Below-average Players')
- 訓練模型, 可針對所需的模型進行屬性挑選
- 切分資料集(test_size=0.33), 並用測試集測試模型
- 分析結果需印出 accuracy、classification report、confusion matrix
- 調整模型, 讓 accuracy 達到 0.9 以上
- 加分題(每大題至多 2.5%)

嘗試使用 matplotlib 等套件將各個演算法結果視覺化

1. Naive Bayes(20%)
2. Decision Trees(20%)
3. Logistic Regression(20%)
4. SVM(20%)

第 5 題

- 取 data_new.csv, 進行 KNN 分析
- 可針對所需的模型進行屬性挑選

5. KNN(20%)

(a) 推薦與 "Neymar Jr" 相像的前五名足球選手

(b) 推薦與 "L. Messi" 相像的前五名足球選手

第 6 題

6. 加分題(10%)

- ✓ 對資料額外進行有趣的分析

- 繳交期限: 6/19 (三) 中午 12:00
- 請轉檔為 PDF 格式, 檔名為: ECT_Project_學號_版本.pdf
- 附上程式碼, 檔名為: ECT_Project_學號_題號.ipynb
- 上傳至 LMS 作業區, 遲交一天扣該次作業成績 5%
- 補交請上傳至 Project 補交區, 僅開放補交至 6/26(三) 中午 12:00