

問題導向之資料科學與機器學習應用 HW1

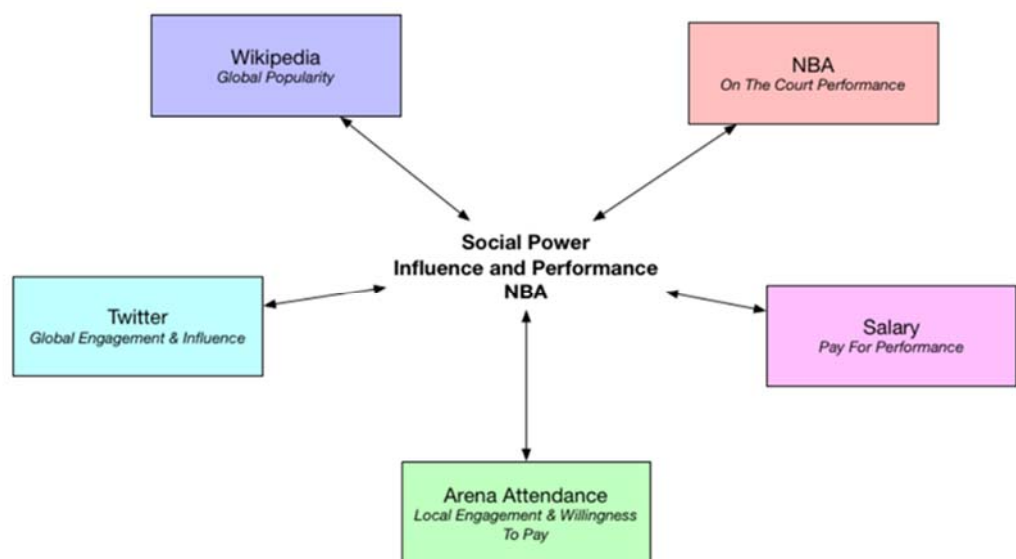
系級:電機 109 姓名:周易廷 學號:F14051091

(A) 資料集:美國職籃 NBA 的社交力量(Social Power NBA)

問題定義:

美國職籃 NBA 近年來開始推行海外政策，讓 NBA 不再只是美國的聯盟，利用資訊的發達與科技的進步讓全世界都參與在這個籃球頂尖殿堂的競技中。而本數據集從幾個面向:球員

Wikipedia 搜尋的流量、球員的 Twitter 用來探討 NBA 球員的社會影響力。



潛在問題:

1. 對於熱衷籃球的球迷來說，爭論哪個球員的能力較強已經成為茶餘飯後的問題，但在這個世界最高籃球殿堂來說，大部分的球員實力都在伯仲之間，每個球員適應的戰術、上場的時間

打的位置都不同，因此很難單純根據得分助攻等比賽數據來比較球員的優劣，由於球迷偏好支持或追蹤能力較佳的球員，因此也可藉由社交面相來探討球員的能力。

2. NBA 是個大型的體育商業聯盟，帶動的經濟效益可達數億美金，而其中很大一部分來自廣告的收益。隨著資訊的發展，現代的媒體已經慢慢轉往網路上，因此也可以藉由數據分析球員的社交影響力，來衡量投注的廣告趨勢，發展新型態的廣告方式，讓收益最佳化。

分析與預測難度：

1. 由於美國職籃有主客場制度，每個球隊都有一個所屬城市，而根據城市大小、人口、市場的不同，可能會影響該隊球員的社交互動量。
2. 球員的薪水高低與球隊勝率相關性不高
3. 一般球迷會偏向與較高技能的球員(明星球員)互動，而功能性球員的社交互動相對少，互動的頻繁程度是依據球員在球場上的表現而非薪水的高低。
4. 代言的收入與球員所屬的球隊勝率有相關性，因此會左右球員轉隊的選擇。

價值：

藉由 NBA 球員的社會影響力，可以來推斷 NBA 球員的能力與所屬該球隊可能勝出的機率、球隊的價值。

(B) 資料：發票或收據資料

價值：

藉由收集發票或收據的資料，來分析特定族群(如學生)的消費習慣與消費能力，可用來預測開發新市場的可行性。