

2. (d) Compare the results in ranking of part (b) and (c), what do you observe?

從結果來看，可以發現B和C的Ranking結果一模一樣。以X-Y平面的單一X對單一Y來解釋。在B小題我們所做的事情就是用一個feature對應一個label，使用linear regression求出一條最符合(X,Y)分布的直線。而在這邊，RMSE越小，代表(X,Y)分布的趨勢越明顯，越符合我們求出的直線。

而C小題叫我們求出每一個feature和label的皮爾森相關係數。皮爾森相關係數用於度量兩個變數X和Y之間的相關程度（線性相依），其值介於-1與1之間。在自然科學領域中，該係數廣泛用於度量兩個變數之間的線性相依程度。這個相關係數也稱作「皮爾森相關係數r」。

由此可知，我們從B小題得到的結論其實和C小題在尋找的東西是一樣的。RMSE越小，代表趨勢越明顯，X與Y的相關性越高。反之，相關係數越高，則趨勢越明顯，RMSE也就越小。