**社群媒體資料分析實務 HW1 報告**

**資工碩一 M10915078 游紹宏**

**演算法原理：Jaccard’s coefficient**

**將輸入之有向圖視為無向圖，計算兩節點間的鄰居數量。**

**加以計算來得到對應係數以評斷兩節點間是否有機會成立連結。**

**在數學定義上，它被用於比較樣本集的相似性的統計量。**

**引申到節點連線預測的觀點：**

**若兩節點間集合的相似度越高，代表兩者間相同的鄰居越多。若相似的鄰居越多，透過三角形連線定理可以推論得到：在相同的鄰居越多的情形下，節點之間會建立節點的機率越高。因此，藉由這個公式並設定判斷條件，即可預測節點是否成立的演算法並運用之。**

**在程式中所使用的方法，透過引入networkx 函式庫。作為處理節點並將之轉換為有向圖的工具，接著引入jaccard\_coefficient( )函式來進行數值計算。若數值大於0.051則判定連結會成立。反之則判定不會成立。**

**經預測之結果再透過csv函示庫輸出成kaggle所需的檔案格式。**