**資安概論108 學年度第 1 學期作業**

分組限制：2~4人

作業目標：

1. 將PTT 至今的文章資料撈取完成並寫入(Elasticsaerch)

* 撈取文章數量多寡關係到最後成績
* 請註明所爬取的文章日期範圍為何
* 請截圖呈現撈取的資料並提供說明包括一共有多少資料以及寫入資料格式型態

1. 針對撈取的資料回答以下問題
   1. 內文提及蔡英文的文章一共有多少
   2. 內文提及韓國瑜的文章一共有多少
   3. 內文提及宋楚瑜的文章有多少
   4. 針對以上問題提出社群媒體對於選舉之看法(300字)
2. 加分題
   1. 為將撈取回來的文字進行斷詞分析並寫入Elasticsearch
   2. 如果有網軍，提出該如何抓出網軍的方式(加分題)

繳交之規則：

* 檔案必須是zip檔案，之中包含兩個：爬蟲程式原始碼(.py或著.ipynb都可以)與pdf報告解釋上述問題。
* 檔案取名之規則是：XX大學資安概論期末作業.zip
* 寄送之mail規則，主旨：XX大學 108 學年度資安概論第1學期作業
* 信件內文須交代，所有組員名字和學號（重要！！）

收作業之mail：openbox008 @cyber00rn.org

**特別注意事項：繳交之規則有誤肯定會關係到給分**

**作業繳交期限為 ：2021/1/3 11:59 截止**

* **不接受補交**