

# 林杰 (Hubert)



清華大學資工系

Python/Java/ C/C++/JS

專長：機器學習

其他：資安、DB、Web

英文能力：TOEIC 870 分

Github：

[hubert0527](https://github.com/hubert0527)

E-mail：

hubert052702@gmail.com

個人網站：

hubert0527.github.io

Kaggle：

goo.gl/o2a7WY

## ➤ 簡介

我是一個自我要求很高、勇於挑戰困難的學生，在轉入資工系後兩年大致完成畢業門檻，並開始修習多門研究所專業課程，並在這些課程中都取得佳績。

於機器學習領域，已修畢三門研究所專業課程，有數項專案實作經驗，並與同學組隊參加 Kaggle 資料科學競賽，於眾多專業資料科學家取得銅牌成績。

目前於微軟 BingGC ( under Cortana ) 團隊中實習，該團隊目標為將機器學習應用在地圖檢索上，建置高效率、高精準度、商業化的分散式系統。

## ➤ 榮譽事蹟

Kaggle competition: Quora Question Pair

Competition : Bronze medal

Kernel sharing : Silver medal

## ➤ 課外紀錄

- 微軟 BingGC ( ML 相關 ) 實習生
- 參與 Dosudo 矽谷讀書會，並共同擔任 editor
- 於 learning by hacking 擔任 editor
- 參與 Taipei Data Science Meetup
- 參與 Nvidia GTC、JSDC conference
- 參與數項 Kaggle data science 競賽

## ➤ 生涯規劃

短期：

希望能更加專注於專業學習與研究上，期許能在專注與努力之下，找到有興趣與值得研究的題目，並將 paper 發表上 top conference。

長期：

希望能持續在深度學習、強化學習等領域中持續鑽研，並得到相關的 research engineer 職位。

## ➤ 相關專案經驗

### Quora Question Pair ( Kaggle competition )

相關連結：[Report](#)、[Github](#)、[Kaggle](#)

我與碩班同學組隊，參加 Kaggle data science 競賽，根據 Quora 提供的資料，辨識兩句問題字串語意相似度。在此比賽中我們獲得銅牌殊榮。

我主要負責 preprocessing、modeling 與部分 feature extraction。使用 word embedding、RNN、CNN、GBDT 等方式建模，實作 ensemble 提升效能，並對 pipeline 進行多項優化，加速訓練時間。

此外，我實作自己設計的 Name entity replacement ( 在投票中得到銀牌 ) 與 Blending 方法 ( 保有 generalizability 並得到額外 1% log loss 提升 )。

### Image Caption

相關連結：[Github](#)、[Kaggle](#)

研究所課程比賽，使用 MS COCO image set 解 captioning 的問題。

使用 word-embedding、CNN、RNN 作為基本框架，實作 beam search、curriculum learning、簡易 attention。過程中我設計自己的演算法，提升 beam search 精準度，讓整體效能額外提升 10%。

在 CIDEr-D 評分下得分 0.58 ( 在 2015 MS-COCO 比賽中 0.4 以上便能躋身前 40 名 )，在這次比賽中得到了第 5 名。

### News Popularity

相關連結：[Github](#)、[Kaggle](#)

研究所課程比賽，根據新聞的 HTML 原始碼預測受歡迎度。

我負責進行 preprocessing、feature 尋找與抽取、model 設計、training、tuning、testing 等所有工作，並在該次競賽中得到第二名的成績。

我能證明自己對機器學習有正確的認識，有卓越的實務能力 ( 獨自完成三人份程式且成果僅次於第一名小組 )，自行實作/應用多項技術並提升效能 ( Boosting、Stacking、CrossValidation、Ensembling )，最終在 private score 的階段驗證 model 具有 generalizability ( public score: 0.572, private score: 0.579 )。

-- Thanks for reading my resume ! --

-- See more projects on Github : [hubert0527](#) --