

林杰 (Hubert)



清華大學資工系大三

C/C++/Python/Java/JS

目前專長：機器學習、前端

學習中：網路安全、DB、後端

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

英文能力：TOEIC 870 分

手機：

0963243221(簡訊佳)

E-mail：

hubert052702@gmail.com

個人網站 (by myself)：

hubert0527.github.io

線上履歷網址：

goo.gl/YCgCBN

➤ 應徵職位

全年實習 (每周 3 至 4 天佳)。

機器學習、資料科學相關為主，後端、軟工相關亦可。

➤ 簡介

目前就讀資工系大三，擅長於實作，至今絕大多數專案都有獨立重現的能力。因偏好修較嚴格的課程而使得 GPA 較低，CS 平均 GPA3.69，等同 A-。

擅於學習是每個人都會自述的特質，但我有諸多實例可以表述自己**特別**擅長，並且也非常努力。其中之一是我在資工領域僅花費兩年半時間，便已近乎達成畢業門檻，並有自信比許多畢業生更有實務能力。

此外，在課餘也非常積極學習，會關注 Kaggle 並參與比賽，自學 Python、JS、前後端、Chrome extension、iOS app、Arduino ...等，並關注尖端技術發展，除了興趣之外，我也認為這些知識與理念能提供不一樣的新觀點與想法。

我的學習範疇不僅限於資工，在小說、圍棋、音樂、棒球上也都有投注非常多時間精力，有許多傑出表現。

在實習中，我希望能獲得更多的「多人協作」、「處理實際問題」、「實作可靠、可信賴的程式」經驗，積極學習之外，也相信自己的能力能回饋公司、實際做出貢獻。

➤ 個人特質

擅長學習新知識與技術

能正視失敗並持續改進

擅於解決問題與思考解決方案

➤ 生涯規劃

短期之內並沒有繼續進修的計畫，將把握大學剩餘一年充實各方面的知識，接著直接進入職場累積經驗、處理真實問題、理解自身不足，隨後才會再次考量是否進修。

長期而言，希望能在不同的經驗中尋找自己的原點、確定想要花費一生鑽研的領域、找到一群一起奮鬥的團隊。

➤ 相關專案經驗

News Popularity (Large-Scale Machine Learning)

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，並有卓越的實務能力 (獨自完成三人份程式且成果僅次於一個小組)，使用 python 自行實作多個演算法 (Boosting、Stacking、Cross Validation、Ensembling)，在不斷失敗中持續思考改進方式，並持續驗證與修正自己的理解。

Image Caption (Deep Learning)

相關連結：[Github](#)、[Kaggle](#)

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

使用 Word-embedding、CNN、RNN 作為基本框架，實作 Attention、Beam Search、Curriculum Learning。過程中我設計新的演算法，提升 Beam Search 精準度，讓整體效能提升 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。經反省後找出問題的癥結，但無奈於硬體限制暫時無法實際驗證。

Data Crawling (MLB game information crawling)

相關連結：[Github](#)

「資料科學概論」課程的一項專案，取得資料是資料科學的重要基礎，我共抓取橫跨 103 年、17 萬場比賽、652 項 feature、總計 389MB、經過清理的 MLB 比賽資訊，並實作 feature 擷取的簡易 API，大幅減輕後續 training phase 的時間壓力。

使用 python 的 request 與 selenium 分別進行一般頁面資料與 JS render 頁面資料的擷取，針對所有可能的例外狀況進行錯誤處理、清理，增加了對於資料擷取與清理的實務經驗。