

黃晨懿

(+886) 986 366 141
✉ chen-yi-huang001@gmail.com
📄 titaneric.github.io/
📍 titaneric
📱 chen-yi-huang

喜歡不間斷學習，熱情的開源貢獻者及盡心盡力的數據科學專長碩士畢業生。擅長資料處理、機器學習相關知識以及資料視覺化。善於使用 Python 語言及有閱讀過大規模軟體專案及追蹤程式碼經驗 (例：CPython, PyTorch & Buddy System in Linux kernel)。

學歷

2018–2020 數據科學碩士，數據科學與工程研究所，國立交通大學。
2014–2018 工程學士，資訊工程學系，元智大學。

工作經驗

Oct 2020–Jan 2021 電腦整合製造部工程師，台積電 新竹，台灣。
○ 負責為廠區內使用者開發 agent 及報表桌面應用程式
○ 開發網路爬蟲下載數百位求職者履歷，並透過自行開發的通用型履歷剖析工具提取個人資料並達成 95% 以上項目正確性
○ 為前人專案撰寫多個可能情況下的整合測試及包含物件關係圖、執行流程等的詳細文件
○ Robotic Process Automation 與報表多國語系支援先驅
July 2018–Jan 2019 網頁全端開發者，國立交通大學資工系計算機中心 新竹，台灣。
○ 為數百位資工系學生開發網頁服務，並主要負責帳號註冊系統
○ 與組員合作討論及設計資料庫 schema 及 API
○ 參與並學習包含測試、部屬、虛擬化環境及 CI/CD 等現代化軟體開發流程

專業能力

數據科學 Pandas, NumPy, PyTorch, 資料前處理, 資料視覺化
網頁開發 Laravel, Vue, MySQL, MongoDB, PHPUnit, Python unittest
系統與工具 Linux, Git, Docker, Gitlab Runner, Travis CI
程式語言 Python (proficient), JavaScript, C, C++, Rust
語言 Chinese (native), English (TOEIC 725)

開源參與

○ 發現眾多深度學習專案中多餘的幕次偏微分計算 @ PyTorch, JAX, Autograd
○ 補足 struct 欄位序列化、反序列化及新增對應測試 @ go-bbs
○ 改善並美化其中一個範例 @ Yew
○ 改寫部分程式使更具 Python 風格 @ Tensorflow
○ Bug 回報 @ Python extension for Visual Studio Code
○ 參與 issue 討論 @ Windows Subsystem for Linux

專案

- Oct 2020 通用型履歷剖析工具.
- 求職候選人個人資料報表給人資及用人主管做參考
 - 開發網路爬蟲下載數百份履歷，定義可能的模式並萃取出需要的個人資料
 - 在工作經驗及學歷的非結構化資料中達成 95 % 資料正確性
- June 2020 即時網路流量異常偵測系統.
- 交大校園行政網路異常流量偵測
 - 設計與實作資料前處理流程
 - 即時處理高達 30 kB/sec 資料流並轉為特徵向量供日後做預測
- Sep 2019 音樂推薦系統.
- 利用 KKBox WSDM 資料做音樂推薦系統
 - 從 Spotify 蒐集額外歌曲相關資料
 - 實作網頁介面視覺化比較推薦歌曲與使用者喜好
- Sep 2019 從頭打造深度學習函式庫.
- 支持自動微分機制的簡單深度學習函示庫
 - 在不同深度學習函示庫中發現問題並提交 PR 解決
 - 允許使用以層的概念建立深度學習模型
- Sep 2018 帳號註冊系統.
- 專為資工系學生設計的帳號註冊系統
 - 以 Laravel MVC 為基礎實作系統附加 Repository 模板
 - 設計資料庫 schema
 - 實作後端商業邏輯以及帳號開通狀態 API

畢業論文

- 標題 以核函數 *attention* 機制解旅行推銷員問題
- 指導教授 蔡錫鈞
- 簡介 研究建立在先前最佳結果的工作上，他們借用 Transformer 解決旅行推銷員問題。因核函數潛在的內積運算，我們將核函數取代原先的 *attention* 機制中的內積運算。我們的實驗結果得到更短的路徑長並僅花費些許時間。

證書及獎項

- Arctic Code Vault Contributor @ GitHub
- Machine Learning with TensorFlow on Google Cloud Platform
- Querying Data with Transact-SQL