陳嘉俊 | 人工智能工程師 | 生成式 AI 與機器學習

香港中環皇后大道中 99 號, 香港, 香港島, Hong Kong
+852 9123 4567 • ka.chun.chan@example.com • https://www.kachun-ai.dev
@kachun-ai • @kachun-ai • @kachun-ai

Basics

- 人工智能工程師，專注生成式 AI、大型語言模型與檢索增強生成（RAG）應用落地
- 具備 5 年以上機器學習與數據工程經驗，熟悉模型訓練、評估、部署及 MLOps 全流程
- 擅長以 Python、PyTorch 及雲端平台建構可擴展的 AI 服務，並推動負責任 AI 實踐
- 具備跨部門協作與技術指導經驗，能將業務需求轉化為可量測的 AI 解決方案

Education

香港科技大學 **https://www.ust.hk**

Master, 計算機科學與工程, Score: 3.9/4.0 *Sep 2018–Jul 2020*

- 主修機器學習與自然語言處理，深入研究大型語言模型與檢索增強生成技術
- 以優異成績完成畢業論文，主題為粵語語音辨識與意圖理解
- 積極參與人工智能實驗室研究項目，累積模型訓練與部署經驗

Courses: 機器學習, 深度學習, 自然語言處理, 分散式系統, 雲端計算, 資料探勘, 演算法設計與分析

香港中文大學 **https://www.cuhk.edu.hk**

Bachelor, 計算機科學, Score: 3.7/4.0 *Sep 2014–Jun 2018*

- 建立扎實的計算機科學基礎，熟悉演算法、資料結構與軟件工程
- 曾參與校內人工智能研究小組，開發影像分類與文字分析專案
- 獲列入學院院長榮譽名單

Courses: 資料結構, 作業系統, 資料庫系統, 機器學習導論, 統計學, 線性代數

Work

智言科技有限公司 **https://www.zhiyan-ai.com**

高級人工智能工程師 *Mar 2022–Present*

- 帶領 4 人團隊設計與落地企業級檢索增強生成（RAG）平台，支援粵語、英語及普通話問答
- 建立大型語言模型評估流程，涵蓋準確度、幻覺率、延遲與成本指標，將模型上線週期縮短 40%
- 開發基於 Python、PyTorch 與 LangChain 的模型服務，並以 Docker 與 Kubernetes 部署至雲端
- 與產品、法務及合規團隊協作，制定負責任 AI 與資料私隱審查規範
- 指導初級工程師與數據科學家，推動程式碼審查、實驗追蹤與 MLOps 最佳實踐

Keywords: 大型語言模型, 檢索增強生成, MLOps, Python, 模型評估, Kubernetes

香港數碼智能有限公司 **https://www.hkdigitalai.com**

機器學習工程師 *Jul 2020–Feb 2022*

- 開發客戶流失預測與推薦系統，運用 XGBoost、LightGBM 及深度學習模型提升轉換率 18%
- 建置自動化特徵工程與模型訓練管道，使用 Airflow、Spark 及 MLflow 管理實驗與版本
- 將模型封裝為 RESTful API，部署於 AWS 並整合實時資料串流，支援每日百萬級預測請求
- 與業務分析師合作定義關鍵指標，透過 A/B 測試驗證模型成效並持續迭代

Keywords: 機器學習, XGBoost, Spark, MLflow, AWS, A/B 測試

Skills

人工智能與機器學習: Expert **Keywords**: 深度學習, 自然語言處理, 大型語言模型, 檢索增強生成, 模型評估, 電腦視覺

程式語言與框架: Advanced **Keywords**: Python, PyTorch, TensorFlow, scikit-learn, LangChain, SQL

資料工程與 MLOps: Advanced **Keywords**: Docker, Kubernetes, MLflow, Airflow, Spark, AWS, GCP