

陳詠欣 | 數據科學家 | 機器學習與大數據分析專家

中國香港, 香港島, 香港, 香港中環皇后大道中 100 號

☎ +852 9876 5432 • ✉ wingyan@wingyanchan.com

🌐 <https://www.wingyanchan.com>

📧 @wingyanchan • 🌐 @wingyanchan • 📺 @wingyanchan

簡介

擁有超過 6 年數據科學及機器學習經驗, 專注於金融科技、智慧城市及零售分析等領域。

- 善於運用統計建模、機器學習及深度學習技術解決實際商業問題
- 熟悉大數據平台 (Spark, Hadoop) 及雲端服務 (AWS, Azure)
- 具備端到端數據產品開發經驗, 由數據收集、特徵工程到模型部署及監控
- 熱衷於把複雜的數據分析結果轉化為清晰的商業決策建議
- 擁有多項專業認證及期刊論文發表經驗

教育背景

香港大學

碩士, 數據科學, 成績: 3.8/4.3

- 以優異成績 (Distinction) 畢業, 專注於機器學習及大數據分析
- 畢業論文題目為「基於深度學習的香港樓價預測模型」, 獲學院最佳論文獎
- 擔任研究助理, 協助教授進行自然語言處理相關研究

課程: 高級機器學習、大數據分析、統計學習理論、深度學習、數據可視化、雲端運算

<https://www.hku.hk/>

2017 年 9 月 - 2019 年 7 月

香港中文大學

學士, 計算機科學, 成績: 3.5/4.3

- 副修統計學, 建立扎實的數學及程式設計基礎
 - 參與多個編程比賽及黑客松, 曾獲校內數據分析比賽冠軍
- 課程: 數據結構、算法設計、數據庫系統、概率與統計、人工智能

<https://www.cuhk.edu.hk/>

2013 年 9 月 - 2017 年 7 月

工作經歷

智庫數據科技有限公司

高級數據科學家

- 領導數據科學團隊, 為銀行及零售客戶開發風險評估及客戶分群模型
- 設計並部署基於 Python 及 Spark 的機器學習平台, 將模型訓練時間縮短 40%
- 與業務部門緊密合作, 將分析結果轉化為可執行的商業策略
- 推動 MLOps 實踐, 建立自動化模型監控及重新訓練流程
- 指導初級數據科學家, 並定期舉辦內部培訓

關鍵字: Python、Spark、MLOps、機器學習、風險建模

香港電訊有限公司

數據科學家

- 分析流動通訊客戶的使用行為及流失率, 建立客戶流失預警模型
- 利用 SQL 及 Hive 處理每日數十億條數據記錄, 製作儀表板供管理層參考
- 開發客戶生命週期價值 (LTV) 模型, 協助市場部制定精準營銷策略

關鍵字: SQL、Hive、客戶流失、Tableau

<https://www.thinktankdata.io>

2020 年 1 月至今

<https://www.hkt.com>

2019 年 8 月 - 2019 年 12 月

語言

粵語: 母語或雙語水平

英語: 完全專業水平

關鍵字: 廣東話

關鍵字: IELTS 8.0

普通話: 完全專業水平

關鍵字: 普通話水平測試二級甲等

專業技能

機器學習: 專家

關鍵字: 監督式學習、非監督式學習、深度學習、模型評估、特徵工程、強化學習

程式語言及工具: 專家

關鍵字: Python、R、SQL、Java、TensorFlow、PyTorch、scikit-learn

大數據平台: 高級

關鍵字: Apache Spark、Hadoop、Hive、Kafka、Airflow、Databricks

雲端及 MLOps: 高級

關鍵字: AWS、Azure、Docker、Kubernetes、MLflow、CI/CD

榮譽

亞太數據科學競賽亞軍

亞太數據科學協會

2021 年 11 月

與團隊成員於 300 多支參賽隊伍中脫穎而出,運用機器學習技術預測城市交通流量,獲亞軍。

最佳畢業論文獎

香港大學統計學系

2019 年 7 月

畢業論文「基於深度學習的香港樓價預測模型」獲評為年度最佳碩士論文。

證書

AWS Certified Machine Learning –Specialty

<https://aws.amazon.com/certification/>

Amazon Web Services

2022 年 3 月

Azure Data Scientist Associate

<https://learn.microsoft.com/en-us/credentials/>

Microsoft

2021 年 8 月

TensorFlow Developer Certificate

<https://www.tensorflow.org/certificate>

Google

2021 年 1 月

出版刊物

IEEE Access

<https://doi.org/10.1109/ACCESS.2022.1234567>

A Hybrid Deep Learning Approach for Urban Traffic Flow Prediction

2022 年 6 月

- 提出結合 CNN 及 LSTM 的混合深度學習模型,用於預測香港市區交通流量
- 實驗結果顯示所提出的模型比傳統 ARIMA 減少 25% 預測誤差
- 文章已獲 IEEE Access 收錄,並被多位學者引用

香港數據科學期刊

<https://www.hkdsj.org/articles/2020/12>

應用自然語言處理於金融文本情感分析

2020 年 12 月

- 探討如何利用 BERT 模型分析上市公司年報及新聞稿中的情感傾向
- 結合金融領域知識詞典,提升情感分析模型於金融文本的準確度

項目

香港樓價預測系統

<https://github.com/wingyanchan/hk-property-price>

一個基於深度學習的香港物業估價及樓價預測平台

2021 年 2 月-2021 年 6 月

- 整合差餉物業估價署、土地註冊處及公共交通數據
 - 使用 XGBoost 及深度神經網絡建立樓價預測模型, R^2 達 0.91
 - 開發 Flask API 及 React 前端,讓用戶輸入物業資料即時獲得估價
- 關鍵字: Python、XGBoost、React、房地產、深度學習

智能客戶服務聊天機器人

<https://github.com/wingyanchan/chatbot>

為銀行客戶服務部門設計的粵語智能聊天機器人

2022 年 9 月-2023 年 3 月

- 利用 RAG(檢索增強生成)架構結合領域知識庫,解答客戶常見問題
- 使用 LangChain 及 OpenAI API 開發,並部署至企業 Kubernetes 集群
- 聊天機器人成功處理 70% 的常見查詢,大幅減輕前線客服人員負擔

關鍵字: RAG、LangChain、Kubernetes、生成式 AI、客戶服務

興趣愛好

人工智能倫理:負責任 AI、數據私隱、公平性

戶外活動:行山、露營、攝影

志願服務

香港數據科學義工網絡

<https://www.datasciencevolunteer.hk>

兼職導師

2021 年 1 月至今

- 為非牟利機構提供免費數據分析諮詢
- 定期主持數據科學工作坊,教授 Python 及數據可視化技巧
- 協助設計「社企數據儀表板」,用於追蹤服務成效