林彥廷 | 資深人工智慧工程師 | 大型語言模型與 MLOps 專家

中國臺灣，台北市，台北市，台北市大安區復興南路一段 390 號，106
+886 912 345 678 • yenting.lin.ai@example.com • https://www.yentinglin.dev
@yentinglin-ai • @yenting-lin-ai • @yentinglin.ai

簡介

- 具備 8 年以上人工智慧與機器學習實務經驗，專精大型語言模型、檢索增強生成與 MLOps。
- 曾領導 5 人跨功能團隊，將 LLM 應用導入企業知識管理、客服與文件摘要流程，提升作業效率 40%。
- 熟悉從資料清洗、模型訓練、評估、部署到監控的端到端流程，並具備雲端與地端混合部署經驗。
- 樂於將技術轉化為可落地的產品價值，擅長跨部門溝通、技術文件撰寫與團隊知識傳承。

教育背景

國立臺灣大學 **https://www.ntu.edu.tw**

碩士，資訊工程學系，成績：4.1/4.3 2015 年 9 月–2017 年 6 月

- 研究主題聚焦於深度學習與自然語言處理，完成以注意力機制為基礎的文字分類研究。
- 擔任課程助教，協助機器學習與資料結構課程之教學與作業批改。
- 與實驗室團隊合作發表 2 篇研討會論文，培養獨立研究與團隊協作能力。

課程：機器學習、深度學習、自然語言處理、電腦視覺、資料探勘、演算法設計與分析、分散式系統、大數據分析

國立清華大學 **https://www.nthu.edu.tw**

學士，資訊工程學系，成績：3.9/4.3 2011 年 9 月–2015 年 6 月

- 奠定紮實的資訊工程基礎，並選修人工智慧與資料探勘相關課程。
- 參與校內專題競賽，以影像辨識應用獲得佳作。

課程：資料結構、演算法、作業系統、資料庫系統、線性代數、機率與統計、計算機網路、程式設計

工作經歷

雲智科技股份有限公司 **https://www.cloudwise-ai.tw**

資深人工智慧工程師 2022 年 1 月至今

- 領導 5 人 AI 團隊，規劃與交付企業級大型語言模型應用，服務超過 20 家企業客戶。
- 設計檢索增強生成（RAG）架構，整合向量資料庫與內部知識庫，將問答準確率提升至 92%。
- 建置 MLOps 流程，使用 MLflow、Docker 與 Kubernetes 管理模型版本、部署與監控，縮短上線時間 50%。
- 建立模型評估與護欄機制，降低幻覺與敏感資訊外洩風險，並制定內部 AI 使用規範。
- 與產品、資料工程及資安團隊協作，將需求轉化為可執行技術方案，並培育 3 位初階工程師。

關鍵字：大型語言模型、檢索增強生成、MLOps、團隊領導、模型評估

睿智數據股份有限公司 **https://www.brightdata-ai.tw**

機器學習工程師 2019 年 3 月–2021 年 12 月

- 開發客戶流失預測與推薦系統，透過特徵工程與集成學習，將預測準確率提升 18%。
- 使用 PyTorch 與 TensorFlow 建立深度學習模型，並以 Airflow 與 Spark 處理每日百萬筆資料。
- 將模型封裝為 RESTful API 並部署至 AWS，支援即時推論與批次預測場景。
- 參與 Agile 開發流程，負責技術文件撰寫、程式碼審查與跨團隊知識分享。

關鍵字：機器學習、PyTorch、Spark、AWS、RESTful API

數字智能股份有限公司 **https://www.digitintelligence.tw**

資料科學家 2017 年 7 月–2019 年 2 月

- 分析行銷與營運數據，建立視覺化儀表板與自動化報表，協助管理層制定決策。
- 運用統計檢定與 A/B 測試評估行銷活動成效，提出優化建議並帶來 12% 轉換率成長。
- 建置資料清理與特徵管線，提升資料品質與分析效率，減少 30% 手動處理時間。
- 與業務與產品團隊密切合作，將數據洞察轉化為具體營運策略。

關鍵字：資料分析、A/B 測試、SQL、視覺化、統計建模

語言

中文：母語或雙語水平　　**關鍵字**：繁體中文、母語

英語：完全專業水平　　**關鍵字**：TOEIC 950、技術文件撰寫、國際團隊協作

專業技能

機器學習與深度學習：專家　　**關鍵字**：Python、PyTorch、TensorFlow、scikit-learn、XGBoost

自然語言處理：高級　　**關鍵字**：Transformers、Hugging Face、LangChain、RAG、Fine-tuning

MLOps 與雲端部署：高級　　**關鍵字**：Docker、Kubernetes、MLflow、AWS SageMaker、GCP Vertex AI

資料工程：高級　　**關鍵字**：SQL、Spark、Airflow、Kafka、BigQuery

程式設計：專家　　**關鍵字**：Python、Go、TypeScript、Shell、C++

榮譽

最佳應用論文獎

台灣人工智慧學會　　2023 年 11 月

- 以「檢索增強生成於企業知識問答之應用」獲得最佳應用論文獎。
- 論文提出混合稀疏與稠密檢索策略，有效提升長文件問答的召回率與準確率。

年度傑出工程師

雲智科技股份有限公司　　2023 年 12 月

- 表彰在大型語言模型產品化與 MLOps 流程建置上的卓越貢獻。
- 帶領團隊完成 3 項企業級 AI 專案，並顯著提升客戶滿意度。

證書

AWS Certified Machine Learning –Specialty　　**https://aws.amazon.com/certification/certified-machine-learning-specialty/**

Amazon Web Services　　2022 年 3 月

NVIDIA Certified Associate –AI Infrastructure　　**https://www.nvidia.com/en-us/training/certification/**

NVIDIA　　2023 年 8 月

Google Cloud Professional Machine Learning Engineer　　**https://cloud.google.com/certification/machine-learning-engineer**

Google Cloud　　2024 年 1 月

出版刊物

2023 台灣人工智慧年會論文集　　**https://www.taiwanai.org.tw/proceedings/2023**

以檢索增強生成提升企業知識問答準確率之研究　　2023 年 11 月

- 提出結合 BM25 與向量檢索的混合式 RAG 架構，改善長尾問題的檢索品質。
- 實驗顯示在企業內部知識庫上，回答準確率較基線模型提升 15%。

中華民國資訊學會通訊　　**https://www.ics.org.tw/journal/2022**

輕量化 Transformer 於邊緣裝置之推論加速　　2022 年 6 月

- 探討知識蒸餾與量化技術在邊緣裝置上的效能與精度權衡。
- 提出適用於資源受限環境的推論優化流程，降低延遲 35%。

項目

企業知識庫問答系統　　**https://github.com/yentinglin-ai/enterprise-rag**

基於檢索增強生成的企業內部知識問答平台　　2023 年 1 月–2023 年 8 月

- 整合 Elasticsearch 與向量資料庫，支援文件上傳、切塊、嵌入與檢索。
- 設計多輪對話與引用來源機制，提升回答可信度與可追溯性。
- 以 FastAPI 與 Docker 部署，支援權限控管與使用量監控。

關鍵字：RAG、LLM、Elasticsearch、FastAPI

智慧客服意圖分類平台 **https://github.com/yentinglin-ai/intent-classifier**

支援多語系與多標籤的客服意圖分類服務 2021 年 3 月–2021 年 12 月

- 使用 BERT 與多標籤分類架構，處理客服來文意圖與緊急程度預測。
- 建立資料標註流程與模型監控指標，持續改善分類表現。
- 將模型部署為微服務，提供即時 API 與批次推論功能。

關鍵字：NLP、BERT、FastAPI、Docker

興趣愛好

開源社群：貢獻套件、技術演講、社群經營

戶外活動：登山、馬拉松、攝影

志願服務

台灣資料科學協會 **https://www.twdsa.org**

志工講師 2021 年 9 月–2023 年 12 月

- 定期開設機器學習入門與 Python 資料分析工作坊，累計指導超過 200 位學員。
- 協助協會規劃年度資料科學年會，負責議程安排與講者邀請。
- 參與偏鄉程式教育計畫，設計適合高中生的 AI 體驗課程。