

## 聯繫方式

shanglin8471@gmail.com

+886-975-601-655

新北市板橋區

https://github.com/ShangLin1606

## 學經歷

## 輔仁大學

金融碩士 • 2021/09 - 2023/06

- 書卷獎
- 專注於機器學習與金融科技領域。

#### 東吳大學

物理學系 • 2013/09 - 2017/06

- 參與中山大學物理交流計畫。
- 與教授共同撰寫多普勒效應研究。

## 技能與工具

• 程式語言: Pytohn、Java

• ML/DL : Scikit-learn, PyTorch

· LLMs: Al Agent, RAG, GraphRAG

• MLOps: MLflow, Dagster, Docker, ELK

• 後端/API: FastAPI, Flask

• 資料庫: PostgreSQL, MongoDB, MySQL

• 資料視覺化: Tableau, PowerBI, Streamlit

• 大數據: Hadoop, Hive, Spark

## 認證

- Big Data 巨量資料分析就業養成班 資展國際(前資策會教研所)
  2024年2月 - 2024年7月
- IBM Data Science Professional Certificate IBM 2025年10月

# 語言

### 英文 - 尚可

# 謝尚霖

## 資料科學家

## 關於我

具備物理與金融雙重背景的 AI 與資料科學專業人士,擅長開發具實際商業價值的AI智能系統。擁有機器學習、LLM 整合(RAG、GraphRAG、AI Agent)與雲端部署的實務經驗,曾成功參與金融科技 AI 平台與法律 AI 系統的建置。熱衷於將資料轉化為可行的洞察與決策依據。

## 工作經驗

2023/11 - 2024/02

### 風險管理人員

聯邦銀行

- 每日監控放款曝險以確保遵循法規。
- 協助機器學習模型的測試與校正,用於風險預測。
- 與資訊部門協作,優化跨部門資料庫邏輯與查詢效率。

2022/09 - 2023/10

## 助教與機器學習營隊講師

輔仁大學金融系

- 規劃機器學習競賽與信用風險個案研討。
- 教授 XGBoost 應用於 P2P 放款違約預測。
- 維護實驗室環境並技術指導學生團隊。

## 專案經驗

#### LegalDocGen - AI 法律文件生成系統(2025/09 至今)

- 使用 GraphRAG 與 LangChain 架構法律文件生成系統。
- 模型 BLEU 分數提升 20%、回應延遲低於 15 秒,實現完整 MLOps 流程。
- 介面採用 Streamlit 與 Plotly 提供互動式操作與視覺化。

#### NeuroQuery - TEXT TO SQL系統(2025/07 - 2025/09)

- 導入「自然語言 → SQL」,不會寫 SQL 也能拿到表格結果與SQL指令。
- 結合 RAG, 理解公司內部術語與欄位別名, 提升複雜報表生成準確度。
- SQL 安全閘預設只讀、強制 LIMIT, 封鎖 DROP/TRUNCATE 等危險指令。
- 提供微調/蒸餾/量化/剪枝腳本,維持效果同時壓低算力與延遲。

### News Summarizer TW - 新聞分類與摘要系統(2025/05 - 2025/07)

- 資料處理:段落切割+摘要策略,結合 LLM 與 Prompt Engineering,處理不同長度文本並維持語意完整性。
- 摘要品質:以 BERTScore 評估,平均 F1 = 0.78,兼顧語意涵蓋與壓縮效率。
- 模型訓練:BGE-m3 語意嵌入+BiLSTM 完成 20 類新聞分類,測試集 F1 = 0.89。

## Grostocks - AI 投資與量化平台 (2024/12 - 2025/05)

- 建構多智能代理架構的 AI 投資平台。
- 以 React 儀表板視覺化市場趨勢。
- 整合 RAG 與 GraphRAG 強化問答與分析能力。
- 自動化 ETL 流程,並透過 FastAPI 與 Docker 部署。

#### 玉山人工智慧挑戰賽專案(2024/11 - 2025/01)

- 任務:長文本檢索優化,聚焦提升語意一致性與檢索精度。
- 方法:設計滑動窗格+語意導向切分,確保 chunk 內涵完整且便於向量化。
- 技術:以 BGE-m3 產生嵌入,配合 FAISS 建索引,強化匹配準確度與查詢速度。

## 美股預測平台(2024/05 - 2024/07)

- 建構 RNN、MLP、SVM 並採用 stacking 模型融合技術。
- 提取 RSI、MACD、Hurst 指數等技術指標作為特徵。
- 使用 cuML 加速訓練流程,效能提升達 4 倍。
- 模型準確率達 80%, 並部署即時預測網站。