

金詩凡

emily00718@gmail.com | 0979527738

<https://www.linkedin.com/in/shi-fan-jin-6a1788117/> | <https://github.com/shifanjin/my-project-job>

基本信息

數據分析師, 具有統計和數學的背景, 扎實的理論功底和邏輯分析能力。擁有豐富的數據處理管理、ETL、數據建模、可視化、優化、分析機器學習模型方面的經驗。擁有團隊管理經驗, 出色的溝通和項目推動能力。熱衷於將分析結果轉化為有價值的信息。

教育經歷

加州大學伯克利分校 伯克利, 美國
統計學專業, 學士 2017.08 – 2019.05

De Anza College 庫柏提諾, 美國
Dean List 榮譽獎 2015 – 2017 2015.08 – 2017.06

相關課程: 統計學、數據分析、機器學習、數據庫系統、時間序列、計算機程序、經濟學、商業分析、定價學

項目經歷

Spam/Ham 郵件分類 2019.04

- 使用機器學習建立數據分類器, 用以區分 Spam 郵件與 Ham 郵件
- 使用 SciKit-Learn Library 和 WordCloud Library 降低過擬合數值已達成優化

Twitter 內容數據分析 2019.03

- 使用 Python 以及 RegEx 轉換處理 twitter 數據
- 以 VADER lexicon 為基礎完成多樣分析, 建立 twitter 數據的情緒檢測

San Francisco 食物安全調查 2019.02

- 使用 Seaborn package 進行數據整合, 分析, 可視化以及建模
- 應用概率抽樣技術完成食物安全評分預測

倫敦股票時間序列預測 2018.11

- 通過對時間序列定價行為和趨勢進行數據分析和轉換來達到英國股票價格預測
- 將統計學因素與交叉驗證和 ACF/PACF 圖等統計技術相結合, 以實現模型的改進
- 使用合適的 ARIMA/SARIMA 模型提高預測精準度

森林火災範圍預測 2018.11

- 使用選定的機器學習模型將精準度提高 7%, 以預測燃燒範圍
- 通過使用線性迴歸, 交叉驗證, 回歸樹, SVM, KNN 等機器學習技術, 達到預測的優化

工作經歷

實習生, 國泰世華銀行, 台北, 台灣 2016.08 – 2016.09

- 幫助客戶開戶, 用 Excel 輸入用戶資料, 調取交易數據
- 整理不動產謄本, 歸納貸款資料

數學和統計助教, De Anza College, 庫柏提諾, 美國 2016.09 – 2017.06

- 舉辦考試複習講座, 並通過簡化數學概念來加深學生印象, 使學生成績提高 20%
- 為在校學生提供數學/統計輔導 (每週 6 小時的個人和小組輔導) 並教導學生用全面性思維去解決問題

知識技能和其他

語言: 英文 (精通)

計算機軟體/語言: R, Python, Java, SQL, C++, Tableau, Excel VBA, Jupyter, Google Analytics

數據/機器學習 Library: NumPy, Pandas, Matplotlib, Seaborn, WordCloud, SciKit-Learn, e1071, Caret, randomForest, glmnet, Survey, Mice, Naniar, ggplot2, dplyr

領導力與組織: 財務部長 (台灣同學會), 書記部長 (Student of Success Club)

興趣愛好: 籃球, 游泳, 音樂作曲編輯 (<https://soundcloud.com/user-54733327>)