

敬啟者:

我是財務金融學系雙主修校學士學程(自訂名稱: 決策資料科學)的柯宥圻，程式年資約為兩年初，最擅長語言是 Python，相當擅長 R, SQL, HTML, CSS，也對 C++, Julia, JS 有一定程度的涉略。目前我在 Python 人工智慧程式設計(統計學課程)擔任助教一職，同時在逸志股份有限公司(豹投資)擔任 Prompt Engineering and finance intern，主要的工作是大型語言模型調研、API 調用，商業模式開發等，希望能夠透過 LLM 微調技術結合股市資料庫，給用戶專屬的投資體驗，以下會詳述。

擅長領域: 文字探勘 NLP, 資料分析, 人工智慧領域, 大型語言模型相關技術。

以下是我的程設相關的課程經歷:

- 110-1 資管系 商管程式設計 A+
- 111-1 資管系 文字探勘初論 A
- 111-2 資管系 大數據與商業分析 A V
- 111-2 資管系 Python 資料分析與機器學習應用 A+
- 111-2 資管系 管理數學 A (大量 Python 應用)
- 111-2 國企系 大數據行銷 A+
- 111-2 工管系 視覺化與現代資料科學 A+
- 111-2 工工所 資料分析方法 A
- 111 全 統計學、計量導論 A/A (R 語言撰寫)
-
- 112-1 電機系 計算機程式暨實習、資料結構、電腦網路概論
- 112-1 資管系 資料庫管理
- 112-1 資工所 雲原生程式應用開發
- 112-1 財金系 財務模型程式應用

有關於程式交易的 Code，存在這裡: 希望可以做到 100 題!

https://github.com/blackwingedkite/TMBA_100strategies_challenge

目前的實習內容：GuTalk AI

由於商業機密因此不能夠多談，但是我們使用了大量 LLM 微調技術、API 串接以及 DB 支援，創造出全新的(至少地球上目前還沒有的)類 ChatGPT 體驗的智能股市助手。我們會透過用戶的問題，分析用戶的投資風格以及習慣，並且隨時改變。每處理一個用戶的輸入問題，我們就借助 LLM 微調的力量為用戶問題進行分群分類、調整、進入 DB 搜尋資料以及統整，也包含圖跟表的製作，最後用戶可以藉由他詢問的問題，得到細緻、有料的答案，加上 3 至 5 張最相關的股市互動性圖表，作為用戶的參考，以及可能的投資建議(但是不能夠太明顯，會違反法律)，最後還有利用 K 棒圖自動分析型態等。如果不是我簽了保密協議，我真的很想要給大家看看我們做的東西。我的職位是 Prompt engineer and finance intern，主要工作內容包含大部分的機器學習、LLM 大型語言模型的調研與開發，還有 LangChain 的流程規劃。

<https://test.gutalk.ai/>

1. 生成式 AI 技術分析用戶問題，結合用戶投資經驗、風險偏好及選股邏輯產出相應動態圖表、文字敘述等相應建議
 - 解決投資人不知道做功課該從何下手的問題
 - 有效做到個人化
2. 收藏投資問題集，每日更新投資建議自動化
 - 透過紀錄用戶平常在研究股票慣用的判斷依據及選股問題，自動更新個人化投資建議報告
 - 設定自選股頁面，產出不同觀點的投資報告 長線、中線、短線以及價值投資的報告書
3. 自動辨識技術型態
 - 辨識頭、肩、頸部、辨識 k 線的 16 種型態
4. 動態追蹤熱門股，把握買進時機
 - 解決以往用戶抱怨在主頁看不到熱門標的的個股分析，導致錯失良機
 - 不知道這些熱門股哪些可以買
5. 選股推薦清單
6. Bar race 圖表功能，掌握短、中長期籌碼量變化
7. 不同回覆語氣及趣味聊天功能，增加趣味度

過往實習內容: AI ChatBOx

https://github.com/blackwingedkite/Alchatbox_forcompany

簡單來說就是使用 Vector Db + LLM API + 前端，幫公司架構了完整的機器人客服，讓用戶可以一進到公司內部就可以馬上跟機器人聊天，獲得資訊，並且推薦公司的產品值得信賴。由於保密的原因也沒辦法將資料全數釋出，但大概就是如 user documentation 那樣。

競賽: 布朗運動、隨機過程相關程式碼。可以利用爬蟲去爬過往股市資料，並且對未來股價進行預測。

https://github.com/blackwingedkite/brownian_motion

110-2 暑假 自主探索學習: 策略分析助手

算是自己的 side project，主要的功能是讓用戶輸入想要偵測的股號、時間段和策略後，可以看到這個策略在過往一段時間的績效狀況。主要的技能是爬蟲、前端介面製作，以及不同策略間的邏輯撰寫與輸出。由於系統是收到輸入訊號才去爬蟲，所以思考速度會稍微慢些，也需要躲避爬蟲的偵測，但其他部分就做得很不錯，也有做到客製化的部分。雖然這裡沒有用到機器學習的技術，全部都是一班的邏輯策略撰寫，但是工作量也很大。有六種策略可以選擇。

https://github.com/blackwingedkite/stock_analysis

https://drive.google.com/file/d/1S_g86s1PDAtQwnQGJVtoNwtl2oKhOLkX/view?usp=sharing

111-2 Python 資料分析與機器學習應用：惡毒語言偵測器

我利用 BERT 模型對資料及進行清洗、前處理以及建模，並製作前端介面，以及連結 ChatGPT 的 API，讓用戶可以慢下來思考:傳送這個惡毒言論前，可以藉由 GPT 的協助，讓你的訊息溫馴一點。主要用到的技術是 NLP、BERT、前端、OpenAI API。此外，我們曾經將這個推上 google 當作插件，但是現在已經停止運作。

https://github.com/blackwingedkite/toxic_comment_detector

介紹簡報與影片:

https://github.com/blackwingedkite/toxic_comment_detector/blob/main/group_K_final_slides.pdf

111-2 大數據與商業分析: 另類商品推薦系統

除了基礎的資料清洗與分析了解客戶的樣貌外，我們也利用 Apriori 演算法及協同過濾，加上用戶的過往登入和瀏覽紀錄，透過他們尋找商品的屬性跟形容詞，推薦以產品個別特色為主要分析對象的產品推薦系統，最後獲得上台報告的機會。

https://github.com/blackwingedkite/BDA2023_final_apriori

介紹簡報與影片:

https://github.com/blackwingedkite/BDA2023_final_apriori/blob/master/BDA_G2_1_Final.pdf

111-2 資料分析方法

這門課是我最喜歡的一門課，主要是學習嚴謹的資料方法，例如多變亮分析、維度縮減等，下半學期則是用更數學的方式介紹各種分類器與深度學習的概念，並利用 python 跟 R 實作。課程中有大量的分析作業，以及期末專案：利用任何好的方式去預測未來一周的每個用戶上線時間，並到 Kaggle 進行競賽。我們利用 CNN+LSTM，以及一些 unsupervised learning 的形式分群，最後得到課堂中第四名。

https://github.com/blackwingedkite/data_analysis

<https://docs.google.com/presentation/d/1hI3jLWptfklxeqKCLiExeqyVAgZndMgiGs2IMWHUd3s/edit?usp=sharing>

111-2 大數據行銷 只有用到 excel 跟 spss 但還是有點關聯。

https://drive.google.com/file/d/1lbDrJvSjX4Sha9gnxVm3WJ0W0Dmzl_Yq/view?usp=sharing

111-2 管理數學 利用 numpy 套件處理線性代數問題 供參

https://github.com/blackwingedkite/Management_Math

111-1 文字探勘初論

作業做的都是文句的分析，以及比較偏實務面的 DA 技巧，而期末專案就是 t
惡毒言論偵測器的簡化版。

https://github.com/blackwingedkite/Text_mining_with_blackmaildetector

110-1 商管程式設計: 管爺散步遊戲

<https://github.com/blackwingedkite/PBC2021-final>

一個簡單的遊戲，使用 PyGame 跟 python 內建的各式前端介面做成的小遊戲。