

Final Project：AI 策略家 — 全域經濟數位孿生與反事實策略模擬

1. AI 的未來能力 (Future Capability)

主題：AI 策略家 (The AI Strategist) — 企業決策的模擬器

能力描述：

目前 AI 雖然擅長分析歷史數據進行預測 (Prediction)，但尚未具備對複雜商業系統進行「動態模擬」與「反事實推演 (Counterfactual reasoning)」的能力。20 年後的 AI 將具備建立「全域經濟數位孿生 (Global Economic Digital Twin)」的能力。

當一位創業者或企業執行長輸入一份商業計畫書（包含定價、產品規格、目標市場）時，AI 不再只是基於歷史數據給出靜態評分，而是能啟動一個由數百萬個「具有高度心理真實性」的 AI 智能體（模擬消費者、競爭對手、供應鏈上下游）組成的虛擬平行宇宙。

應用場景與意義：

- **創業壓力測試**：創業者在投入真實資金與時間之前，可以在虛擬世界中運行 10,000 次模擬，測試在不同市場衝擊（如金融海嘯、對手發動價格戰、原物料上漲）下的存活率。
- **政策制定模擬**：政府可用於模擬新稅收政策或補貼對中小企業生存率的連鎖反應。

這項能力將商業決策從「基於直覺與歷史歸納的賭博」轉變為「基於大規模模擬的科學實驗」，能大幅降低創業失敗率，並優化社會資源配置。

2. 所需的成分與資源 (Ingredients)

要實現上述「AI 策略家」的宏大願景，需要以下關鍵成分的深度整合：

- **資料 (Data)：異質性動態數據與行為圖譜**
 - 需要超越財務報表的多模態即時數據，包含社群媒體情緒、供應鏈 IoT 數據、人口普查細節，用來構建逼真的「虛擬環境」。
 - **人類行為圖譜 (Human Behavioral Graphs)**：需要海量的人類決策數據（點擊流量、消費紀錄、談判對話），用來訓練 AI 智能體的「心理模型」，使其能展現非理性、從眾、品牌忠誠等

真實人類行為。

- **工具 (Tools)：LLM Agents 與多智能體強化學習**

- **大型語言模型 (LLMs) 與認知架構**：作為智能體的「大腦」，賦予它們理解敘事、產生信念與進行複雜推理的能力，而非僅僅是簡單的規則反應。
- **多智能體強化學習 (Multi-Agent Reinforcement Learning, MARL)**：處理數百萬智能體之間的動態博弈。競爭對手會學習你的策略並反擊，這需要 MARL 來模擬策略的演化。
- **因果推論 (Causal Inference)**：區分「相關性」與「因果性」，確保模擬結果的邏輯正確性（例如：是降價導致營收增加，還是季節因素？）。

- **硬體與環境 (Hardware / Environment)**

- **超大規模並行模擬器**：需要能同時運行數百萬個異質智能體互動的運算架構（類似 Matrix 的伺服器），支援極低延遲的狀態更新。

- **學習架構 (Learning Setup)**

- **Sim-to-Real Transfer (虛實遷移)**：模型需在虛擬環境訓練，並不斷利用真實世界的反饋（Real-world feedback）來校準其模擬參數，縮小「模擬落差 (Simulation Gap)」。

3. 涉及的機器學習類型 (Type of ML)

- **主要類型：強化學習 (Reinforcement Learning) 與生成式模型 (Generative Models) 的結合**

- **理由：**

- 動態博弈 (Why RL?)**：商業競爭本質上是一個多方互動的過程。當我方改變策略（例如降價），競爭對手會做出反應（例如跟進或打廣告），消費者偏好也會隨之改變。監督式學習無法處理這種「策略互動」與「延遲獎勵」，只有強化學習能讓智能體在互動中學習最佳策略。
- 創造未來 (Why Generative?)**：我們的目標不是分類或回歸歷史數據，而是「生成」一個從未發生過的未來（反事實場景）。整個市場模擬本身就是一個巨大的生成過程。

- **資料來源與目標訊號：**

- **資料來源 (State)**：模擬環境中的當前狀態（市場價格分佈、消費者滿意度、對手庫存）。
- **目標訊號 (Reward)**：
 - 對企業智能體：長期利潤 (Long-term Profit) 或 市場佔有率 (Market Share)。
 - 對消費者智能體：個人效用最大化 (Utility Maximization) 或 心理滿足感。

4. 第一步的「可實作模型問題 (Solvable Model Problem)」

為了邁向 20 年後的宏大願景，設計並實作了一個簡化版的模型問題，剝離掉複雜的 LLM 心理學，專注於「市場機制」與「策略最佳化」的核心邏輯。

4.1 問題設計

- **題目：「Sim-Coffee」：基於智能體的咖啡店選址與定價模擬器**
- **對應關係**：此問題捕捉了最終目標的核心運作循環：「策略輸入 → 多智能體互動 → 湧現經濟結果」。
- **輸入 (Input)**：創業者設定的策略向量
[定價 (Price), 品質 (Quality), 地點 (Location_X, Location_Y)]。
- **輸出 (Output)**：
 - i. 模擬 365 天的每日利潤曲線。
 - ii. 全域搜索後的最佳策略組合與預期年利潤。
 - iii. 利潤熱力圖 (Heatmap) 與市場競爭地圖。
- **任務目標**：在一個已有競爭者存在的網格地圖中，找出能最大化年利潤的開店策略。

4.2 模型與方法

- **選擇模型**：基於智能體的模型 (Agent-Based Model, ABM)。
- **理由**：
 - 市場宏觀結果（營收）是由大量個體（消費者）的微觀決策所「湧現 (Emerge)」出來的。
 - 傳統迴歸模型難以捕捉「消費者分眾」（小資族 vs. 菁英族）以及「非線性成本結構」帶來的突變點。
- **核心機制設計**：
 - i. **消費者決策 (Utility Function)**：消費者根據價格、品質感受（邊際效用遞減）、距離來計算效用。
$$Utility = (w_q \cdot \sqrt{Quality}) - (w_p \cdot Price) - (w_d \cdot Distance)$$
 - ii. **市場分眾 (Segmentation)**：引入「預算限制 (Budget Constraint)」與「品質門檻 (Quality Threshold)」。
 - **小資族**：預算低，對價格敏感。
 - **菁英族**：預算高，但有最低品質要求（未達一定品質不消費）。
 - iii. **非線性供給端 (Non-linear Supply)**：成本隨品質呈指數成長 ($Cost \propto Quality^2$)，模擬真實世界追求極致品質的高昂代價。

4.3 實作與結果

使用 Python 實作了 Consumer、CoffeeShop 與 MarketSimulator 類別。透過蒙地卡羅方法 (Monte Carlo Sampling) 進行全域策略搜索。

結果展示與分析：

1. 策略對決 (營收曲線)：

比較了「高品質高價」與「低品質低價」兩種策略。在考慮了菁英族的品質門檻後，高品質策略雖然客群較少，但因單價高且競爭者（被品質門檻過濾掉）少，反而能產生穩定的高利潤。

2. 利潤熱力圖 (Profit Heatmap)：

熱力圖顯示了明顯的「死亡區 (Death Zone)」(左上角：高品質低價格，導致成本高於售價) 以及「策略甜蜜點 (Sweet Spot)」。結果顯示，最佳解並非單純的「越便宜越好」，而是在成本曲線與消費者願付價格之間取得平衡。

3. 全域最佳化地圖：

透過全域搜索，模型在二維地圖上標示出了「WINNER」的最佳開店位置。

- **發現：**最佳位置通常位於競爭對手的空隙（服務被忽略的客群），或是緊鄰競爭者但提供更優的性價比（搶客）。
- **數據：**例如在某次模擬中，冠軍策略是定價 \$43 / 品質 1，這顯示在大量小資族存在的市場環境下，薄利多銷的策略擊敗了精品策略。

4.4 討論

• 從這個簡化問題中學到了什麼？

- 結構決定策略：**發現單純調整權重無法改變「低價為王」的局面。必須引入「硬門檻」（如預算上限、品質底線）和「非線性成本」，才能模擬出真實世界中「精品」與「平價」共存的市場樣貌。
- 湧現現象：**微小的參數調整（例如提高一點品質成本係數），會導致整個市場的最佳策略從「高品質」瞬間崩塌為「低品質」，這揭示了商業環境的脆弱性與非連續性。

• 它揭示了未來要解決的大問題中，哪個面向的困難或關鍵？

這個 Toy Model 最大的簡化在於「智能體的靜態性」。

- 行為真實性：**目前的消費者是基於公式的機器人。未來的挑戰在於如何結合 LLM，讓 Agent 具備「被行銷影響」、「產生品牌信仰」或「非理性從眾」的能力。
- 動態博弈：**目前的競爭對手是固定的。在 20 年的願景中，對手應該會觀察我們的定價並發動反擊 (Price War)。解決**多智能體之間的動態均衡 (Nash Equilibrium in dynamic settings)**將是技術上最大的挑戰。