

國立臺中教育大學資訊工程學系
113 級畢業專題報告

人工智慧在不完美信息的應用
專案編號：NTCU-CS-PRJ-113-04

組員：ADT109111 蔡羽玟

ACS108801 林靜雪

ACS109150 林裕晉

指導老師：李宜軒 教授

中華民國 112 年 12 月

摘 要

近十年來 AI 日益漸進，不斷挑戰突破人腦認知，在各類遊戲中接連戰勝人類玩家選手，以德州撲克來說，卡耐基梅隆大學開發的人工智慧系統 Libratus 在比賽中戰勝四位德州撲克頂級選手，獲得最終勝利。Facebook(今 Meta)也於幾年前，與卡內基美隆大學合作發展 AI 機器人 Pluribus，已在無限注賽局中打敗幾名德州撲克專家。

隨著限注的雙人德州撲克對局已被 AI 破解，我們想延續前幾屆的精神將德州撲克在人工智慧的應用上繼續拓展至多人對局的方面。假如有多於兩者玩家，情況就更為複雜，因為其中一名對手的失誤，可能令另一名對手得益，更符合充滿隨機性的現實世界，並期望能提高 AI 勝率，因此我們新增了選擇對局人數的功能，在遊戲開始之前能讓玩家選擇一對二或一對三的牌局，增加遊戲與玩家間的互動，並改良遊戲體驗。此外，我們也預計將不同的 AI 同時導入牌局期待能藉此展示出在同一個牌局中，不同 AI 決策會有什麼不同的影響。

本專題採用的資料格式和上一屆同樣為二維編碼，不同之處在於新增了動作順序，也更改了牌型強度、手牌等級的判斷標準，並將上一屆之 CNN 演算法重新建構，以達到更佳的判斷結果。期望能提升 AI 在德州撲克上的勝率，使用 CNN 演算法訓練出卓越的 AI 選手。

在資料傳輸方面，我們改用 websocket 技術。有別於傳統 AJAX 需要不斷刷新頁面，在 websocket API 中，瀏覽器和服務器只需要完成一次握手，兩者之間就直接可以創建持久性的連接，並進行雙向數據傳輸，能有效節省寬頻及伺服器資源。

目 錄

| | |
|-----------------------------|-----|
| 摘要 | i |
| 目錄 | ii |
| 表目錄 | iii |
| 圖目錄 | iv |
| 一、前言 | 1 |
| 1-1 動機 | 1 |
| 1-2 專題沿革與目標 | 2 |
| 二、歷屆專題與相關概念研究 | 3 |
| 2-1 歷屆專題 | 3 |
| 2-2 相關概念研究 | 3 |
| 三、專題內容 | 7 |
| 3-1 資料處理 | 7 |
| 3-2 演算法 | 9 |
| 3-3 判定 All_in 與 Raise | 10 |
| 四、專題成果 | 11 |
| 4-1 真人 v.s AI | 11 |
| 4-2 AI v.s AI | 11 |
| 4-3 自動生成訓練資料 | 12 |
| 4-4 數據結果 | 13 |
| 五、未來展望 | 15 |
| 5-1 往強化式學習導向發展 | 15 |
| 5-2 採用新現有資料集 | 15 |
| 5-3 生成新資料集 | 15 |
| 5-4 發一個多玩家對局模式 | 15 |
| 參考文獻 | 16 |

表 目 錄

| | |
|----------------------|---|
| 表 1 專題沿革 | 2 |
| 表 2 不完美訊息與完美訊息 | 4 |
| 表 3 德州撲克牌型組合 | 5 |
| 表 4 對局資料類別及局數 | 7 |

圖 目 錄

| | |
|--|----|
| 圖 1 AJAX polling 與 websocket 示意圖 | 6 |
| 圖 2 本屆二維陣列資料欄位 | 7 |
| 圖 3 本屆手牌等級判斷表 | 8 |
| 圖 4 本屆二維陣列資料輸出形式..... | 8 |
| 圖 5 (a) 上屆混淆矩陣 | 9 |
| 圖 5 (b) 本屆混淆矩陣 | 9 |
| 圖 6 損失率曲線 | 10 |
| 圖 7 All_in 與 Raise 流程圖 | 10 |
| 圖 8 真人對 AI 介面示意圖 | 11 |
| 圖 9 AI 對 AI 介面示意圖 | 12 |
| 圖 10 (a) AI 對 AI 數據結果 | 13 |
| 圖 10 (b) AI 對 AI 數據結果 | 14 |
| 圖 10 (c) AI 對 AI 數據結果 | 14 |

一、前言

1-1 動機

近十年來 AI 日益漸進，不斷挑戰突破人腦認知，在各類遊戲中接連戰勝人類玩家選手，以德州撲克來說，卡耐基梅隆大學開發的人工智慧系統 Libratus 在比賽中戰勝四位德州撲克頂級選手，獲得最終勝利。Facebook(今 Meta)也於幾年前，與卡內基美隆大學合作發展 AI 機器人 Pluribus，已在無限注賽局中打敗幾名德州撲克專家。

隨著限注的雙人德州撲克對局已被 AI 破解，我們想延續前幾屆的精神將德州撲克在人工智慧的應用上繼續拓展至多人對局的方面。假如有多於兩者玩家，情況就更為複雜，因為其中一名對手的失誤，可能另一名對手得益，更符合充滿隨機性的現實世界，並期望能提高 AI 勝率，因此我們新增了選擇對局人數的功能，在遊戲開始之前能讓玩家選擇一對二或一對三的牌局，此外，我們也預計將不同的 AI 同時導入牌局，例如玩家可與選擇與 CNN、RF 等，進行對戰，藉此增加遊戲與玩家間的互動，並展示出在同一個牌局中，不同 AI 決策會有甚麼不同的影響，同時改良遊戲體驗。

本專題採用的資料格式和上一屆同樣為二維編碼，不同之處在於新增了動作順序，也更改了牌型強度、手牌等級的判斷標準，輸出後每局資料數為 $[12 \times 13(\text{行}) \times \text{動作數}(\text{列})]$ ，並延用上一屆之 CNN 演算法，期待能提升 AI 勝率。

在資料傳輸方面，我們將資料傳輸方式改為 websocket 技術，前幾屆利用 AJAX 輪詢進行客戶端和伺服器端的資料傳輸。輪詢是在特定的時間間隔，由瀏覽器對伺服器發出 HTTP 請求，然後由伺服器返回最新的數據給客戶端的瀏覽器 websocket 作為在 html5 後引入的功能。有別於傳統 AJAX 需要不斷刷新頁面，websocket 能夠讓瀏覽器和伺服器以直接建立 socket 的方式更新瀏覽器顯示的資料，在 websocket API 中，瀏覽器和服務器只需要完成一次握手，兩者之間就直接可以創建持久性的連接，並進行雙向數據傳輸，能有效節省寬頻及伺服器資源。

1-2 專題沿革與目標

本專題為第四屆，如表 1 所示，起初從第一屆的決策樹、第二屆的隨機森林，發展至第三屆選擇 CNN 作為新的演算法。雖然我們依然採用了 CNN 模型，但進行了大量的改變，除了 CNN 模型架構的修改外，我們還改變了原始的訓練資料格式，我們的目標不僅是延續前幾屆專題提升 AI 在對局中的勝率，而且將前幾屆的 AI 模型（例如決策樹、隨機森林）與本屆的新 CNN 模型結合起來，進行多 AI 的多人對局。最終目標是大幅提升上屆 CNN 模型的準確率，也希望透過記錄多 AI 對局資料來產生新的訓練資料。

表 1. 專題沿革

| 第一屆－決策樹 | 第二屆－隨機森林 |
|--|--|
| <ol style="list-style-type: none"> 1. 實際對局資料蒐集、一對一有秀手牌 2. 決策樹模型訓練 AI 3. 使用者操作介面 UI 4. 預測正確率約 50% | <ol style="list-style-type: none"> 1. AI 模型使用隨機森林取代決策樹 2. 提高訓練及測試資料數量、分割多人牌局 3. 提高 AI 模型預測的正確率 60% 4. 進行 AI 與人對戰 |
| 第三屆－CNN | 第四屆(本屆)－CNN |
| <ol style="list-style-type: none"> 1. 更動訓練 / 測試資料格式 2. 提高預測正確率 3. 提高 AI 與人對戰勝率 4. 增加隨機森林與本屆 CNN 對決 | <ol style="list-style-type: none"> 1. 節省寬頻及伺服器資源 2. 提高預測正確率 3. 改為多人牌局（1 個玩家對 1-3 個 AI） 4. 變更資料編碼格式 5. 加入前幾屆 AI 模型共同對局 |

二、 歷屆專題與相關概念研究

2-1 歷屆專題

- 決策樹

第一屆的決策樹會根據訓練資料產生一棵樹，依據訓練出來的規則來對新樣本進行預測。決策樹演算法可以使用不同的方式來評估分枝的好壞(亂度)，例如 Information gain、Gain ratio、Gini index。依據訓練資料找出合適的規則，最終生成一個規則樹來決策所有事情，其目的使每一個決策能夠使訊息增益最大化。在資料格式採用的是一維陣列。

- 隨機森林

第二屆的隨機森林基本原理是，結合多顆 CART 樹 (CART 樹為使用 GINI 算法的決策樹)，並加入隨機分配的訓練資料，以大幅增進最終的運算結果。顧名思義就是由許多不同的決策樹所組成的一個學習器，其想法就是結合多個「弱學習器」來建構一個更強的模型。在資料格式採用的亦是使用一維陣列。

- CNN

第三屆的 CNN 採用 AlexNet 模型。一般的 CNN 模型是一或多個卷積層和 Pool 層組成。卷積層透過相同的幾何轉換應用，在不同空間和區塊擷取空間的局部 Pattern；Pool 層允許在空間中隨著特徵數量增加來處理資料。卷積神經網路通常以 Flatten 層或全局 Pool 層作為盡頭，越多卷積層越能從低級特徵中迭代取出更複雜的特徵，進行最後分類或迴歸。

當屆使用的 AlexNet 模型因 Pool 層的特性是放大擷取訓練資料之特徵，而此特性將導致牌局完整性被破壞，且因需要完整之牌局資訊，因此將 Pool 層移除。

2-2 相關概念研究

- 不完美訊息與完美訊息

完美訊息遊戲 (Perfect information game) 以及不完美遊戲訊息遊戲 (Imperfect information game) 為兩種完全相反的遊戲概念。完美訊息遊戲指在遊戲中雙方玩家的

行動順序為固定，後手行動的玩家可以觀測到先手玩家的行動，並以此為基礎，後手玩家藉由先手玩家的行動以及自身的行動來預測先手玩家的下一步動作，而遊戲中所有的資訊對雙方完全公開，雙方所擁有的資訊相同，換句話說遊戲雙方並不存在資訊不對稱之情形。例如：圍棋、西洋棋、象棋等大多數棋類遊戲皆屬於完美訊息遊戲。

不完美訊息遊戲則泛指無法觀測別人行動訊息的遊戲，遊戲中的資訊對所有玩家並非完全公開，因此為每位玩家所持有的資訊並不相同，因而造成資訊盲區。且由於資訊盲區的存在使得即使可以觀測到其他玩家的行動，導致無法預測玩家行動。例如：麻將、德州撲克及大部分的撲克遊戲皆屬於此類。

除此之外，完美訊息遊戲還需要符合以下兩個條件：1. 玩家間的行動需要存在確切的先後順序；2. 可以觀察行動行為。只要兩個條件其中之一不符合，這個遊戲便屬於不完美訊息博弈。由於上述條件的存在，不完美訊息遊戲較為符合現實世界發生的情形。

完美訊息與不完美訊息遊戲兩者間的不同之處，如表 2 所示，主要在於順序、資訊透明度和對手行動是否完全公開此三面向。由此三面向來看，德州撲克的遊戲過程與圍棋、象棋等截然不同，論順序，玩家並非每輪皆有出牌，論資訊透明度與對手行動，德州撲克在遊戲過程中也並非完全公開牌面及出牌動作，故將德州撲克歸類為不完美訊息遊戲。

表 2. 不完美訊息與完美訊息

| | 完美訊息遊戲 | 不完美訊息遊戲 |
|--------|-------------------|--------------------|
| 順序 | 有固定的先後順序 | 不一定有固定順序 |
| 資訊透明程度 | 遊戲資訊完全公開透明 | 遊戲資訊不完全、公開透明 |
| 對手行動 | 全數公開 | 部份公開 |
| 實際例子 | 圍棋、西洋棋、象棋 | 麻將、德州撲克 |
| 遊戲 AI | Deep Blue、AlphaGo | Libratus、DeepStack |

· 德州撲克規則

德州撲克使用一般撲克牌進行遊戲，包含 4 種花色牌，每種花色牌有 1 到 13 點花色牌，共 52 張。德州每位玩家初始擁有 2 張私人手牌，牌桌中央會放置公共牌，在發牌並達成第一次開牌條件後，放置 3 張公共牌，開始由各玩家輪流決定是否加注。在達成開牌條件後，放置第 4 張公共牌在所有玩家執行動作且達成開牌條件，放置最後一張公共牌。在此輪行動達成開牌條件，所有玩家公開各自手牌，由各玩家的 2 張手牌，加上牌桌上的 5 張公共牌做出最大組合，比較牌型大小決定勝負。

如下方表 3 所示，牌型組合共有 10 種，牌型由小到大分別為高牌、一對(2 張牌點數相同)、兩對(2 張牌點數相同共 2 組)、三條(3 張牌點數相同加 2 張散牌)、順子(5 張牌點數連續但不限花色)、同花(5 張牌花色相同但點數不限)、葫蘆(3 張牌點數相同加 2 張牌點數相同)、四條(或稱鐵支 4 張點數相同加 1 張散牌)、同花順(5 張牌點數連續且花色相同)。

表 3. 德州撲克牌型組合

| 名稱 | 牌型組合 | 名稱 | 牌型組合 |
|---------------------------------|---|-------------------------|---|
| 皇家同花順 (Royal Straight Flush) |  | 同花順 (Straight Flush) |  |
| 四條 (Four of a kind) |  | 葫蘆 (Full house) |  |
| 同花 (Flush) |  | 順子 (Straight) |  |
| 三條 (Three of a kind) |  | 兩對 (Two pair) |  |
| 一對 (Pair) |  | 高牌 (High card) |  |

· 資料傳輸

在資料傳輸方面，我們將資料傳輸方式改為 websocket 技術，前幾屆利用 AJAX 輪詢進行客戶端和伺服器端的資料傳輸。輪詢是在特定的時間間隔，由瀏覽器對伺服器發出 HTTP 請求，然後由伺服器返回最新的數據給客戶端的瀏覽器。websocket 作為在 html5 後引入的功能。

如圖 1 所示，有別於傳統 AJAX 需要不斷刷新頁面，websocket 能夠讓瀏覽器和伺服器以直接建立 socket 的方式更新瀏覽器顯示的資料，在 websocket API 中，瀏覽器和服務器只需要完成一次握手，兩者之間就直接可以創建持久性的連接，並進行雙向數據傳輸，能有效節省寬頻及伺服器資源。

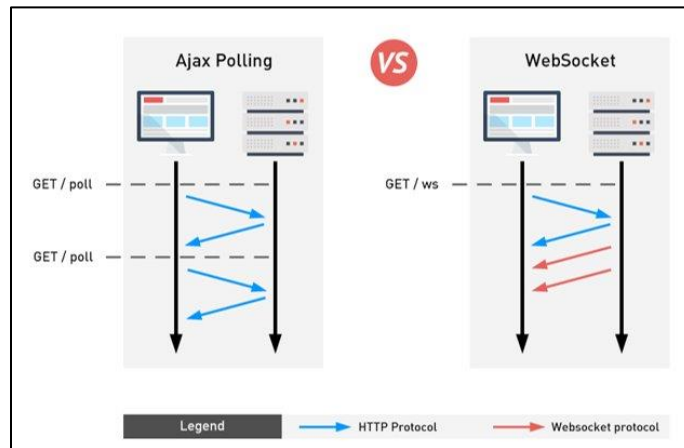


圖 1. AJAX polling 與 websocket 示意圖

三、 專題內容

3-1 資料處理

- 訓練資料來源： Outflopped
- 對局資料：採用三人、四人有秀手牌對局，如下表 4 所示。

表 4. 對局資料類別及局數

| | 全部原始資料(9,283,500 局) | | |
|-------|---------------------|----------|----------|
| 資料分類 | 一對一有秀手牌（上屆） | 三人有秀手牌對局 | 四人有秀手牌對局 |
| 局數（局） | 159,691 | 37,642 | 120,326 |

- 資料格式：

如圖 2 所示，我們採用了與上一屆相同的二維編碼資料格式，但針對多個方面進行了更新。新的資料格式不僅新增了動作順序的記錄，還改變了牌型強度、手牌等級（如圖 3 所示）的判定標準，並對籌碼與大小盲的記錄格式進行了調整。每個動作都被儲存為一個 12x13 的矩陣，並延用上一屆之 CNN 演算法再進行調整，使模型能更準確的預測玩家的決策。

| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | J | Q | K | A | C h i p s 排名 庄家 |
|---|----|----|----|----|----|----|----|----|----|----|----|----|----|-----------------------------------|
| Strength | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | |
| Flop0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| Flop1 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| Flop2 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| Turn | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | |
| River | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | |
| Hand0 | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| Hand1 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | |
| Hands_level | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | |
| chips | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| Action | 0 | 1 | 2 | 8 | 3 | 4 | -1 | -1 | -1 | -1 | -1 | -1 | -1 | |
| B_S | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| player1 player2 player3 player4 Blind/show | | | | | | | | | | | | | | |

圖 2. 本屆二維陣列資料欄位

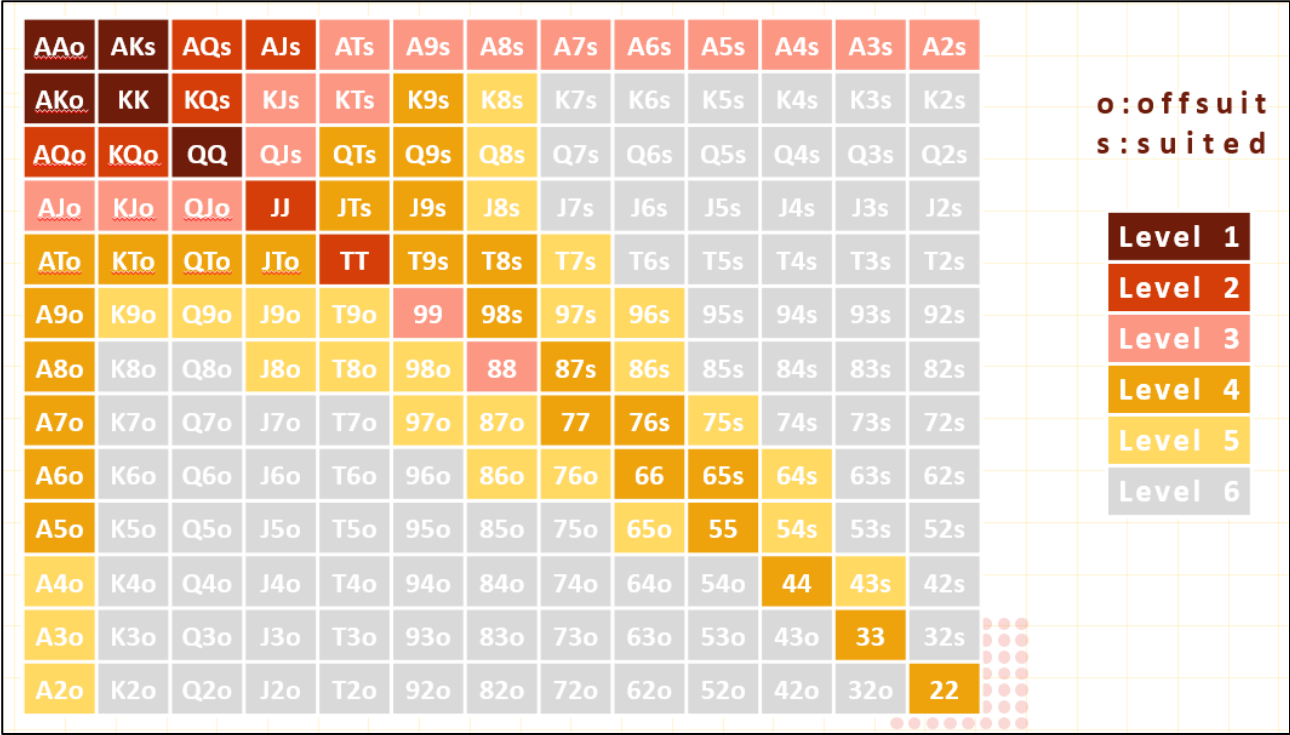


圖 3. 本屆手牌等級判斷表

資料轉換：

在資料轉換的初始階段，我們將原始文字資料轉換成了一維編碼格式，接著進一步轉換成符合 CNN 輸入格式的二維矩陣，如圖 4 所示。

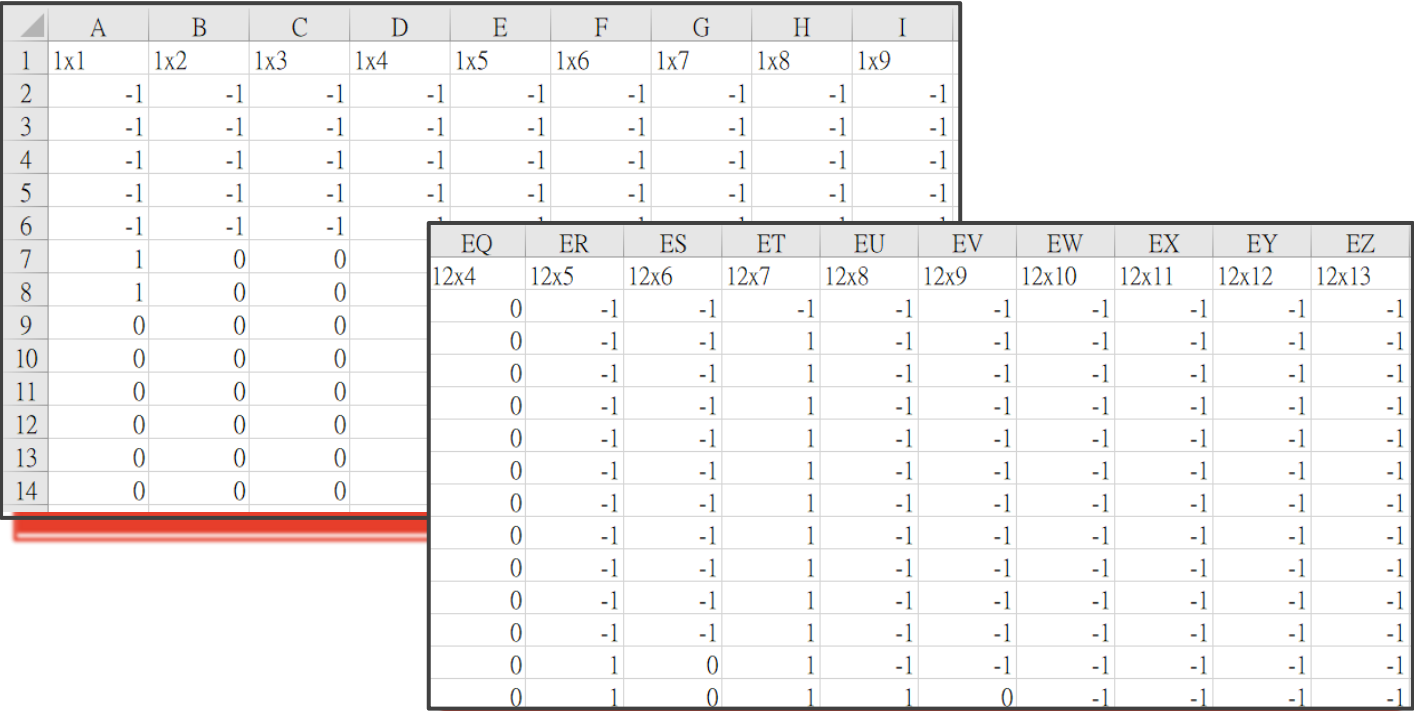


圖 4. 本屆二維陣列資料輸出形式

3-2 演算法

上一屆的 CNN 模型中由於採用了 AlexNet 模型，而其中池化層因為採用的最大池化(Maxpooling)技術會破壞資料完整性而在模型中被捨去。在本屆執行多人對局的 CNN 模型訓練時發現，過往 CNN 的 AlexNet 模型過於複雜對於多人對局的訓練並無幫助，於是我們決定自行建立一個較為簡單的 CNN 模型。該模型是由四個捲積層與三個池化層組成，相較於原先的 AlexNet 模型本屆的模型更能快速且準確的得到結果。

在進行模組調整與訓練時，藉由訓練後產生的損失曲線我們發現本屆的模組在訓練時的損失值大幅擺盪且略有過擬合的發生，為了解決這個現象我們使用了學習率衰減(Learning Rate Decay)策略，藉由定期衰減的方式在每 10 個 epochs 後學習率會以 0.8 的衰減倍率進行更新，也就是每次衰減後的學習率是前一輪的 80%。這樣的學習率衰減策略有助於控制模型在訓練過程中的學習速度，避免學習率過大導致不穩定的訓練，同時也有助於模型更好地收斂到最佳參數配置。

相較於上屆 CNN，本屆在準確率及損失率皆有獲得大幅的提升。在整體準確率的部分由原先的 0.61 提升到 0.8，而損失率則是由原先的 0.88 降至 0.45，如圖 6 所示。可以看到在跟注(call)的部分本屆有顯著的提升，然而在加注(raise)及下注(bet)依然無法獲得明顯提升，我們認為原因是加注及下注在訓練資料的部分加注及下注的比例較少，導致 CNN 模型無法做出良好的判斷。在圖 5 中可以看到本屆的 CNN 模型在套用學習率衰減策略後，損失率能趨於平緩讓模型能順利收斂。

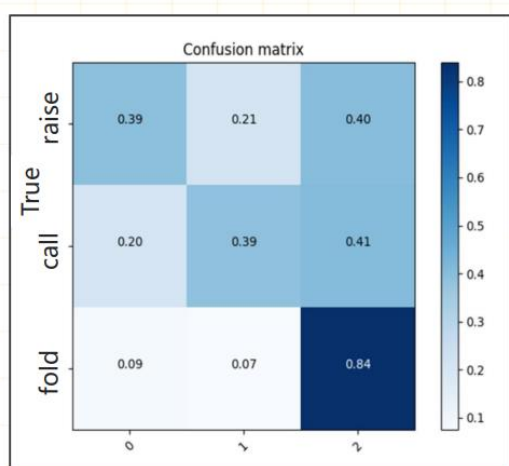


圖 5 (a). 上屆混淆矩陣

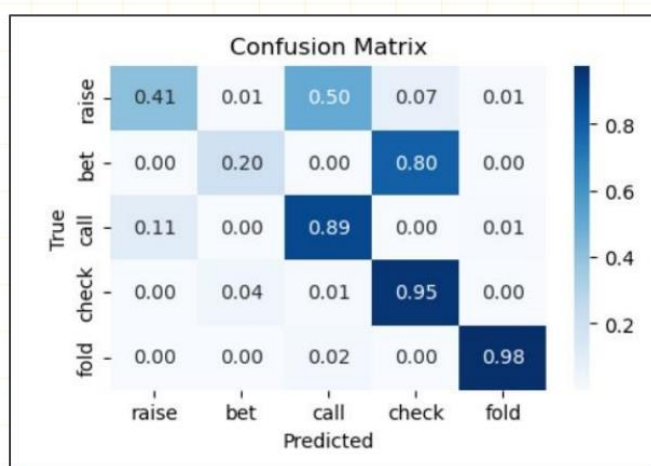


圖 5 (a).本屆混淆矩陣



圖 6. 損失率曲線

3-3 自行判定 All_in 與 Raise 金額之流程

由於 All_in 在訓練資料中的筆數嚴重不足導致模型無法對於 All_in 做出準確的預測，為了改善這個狀況我們決定將 All_in 這個 label 合併入 Raise 的 label 中。圖 7 為 All_in 與 Raise 的流程圖，在模型預測 Raise 後會透過外部程式依據目前手牌的強度 strength 決定是否 All_in 或加注籌碼數量，若是手牌強度 strength 達到一定的標準就執行 All_in 否則依強度區段選擇相對應的籌碼數加注。

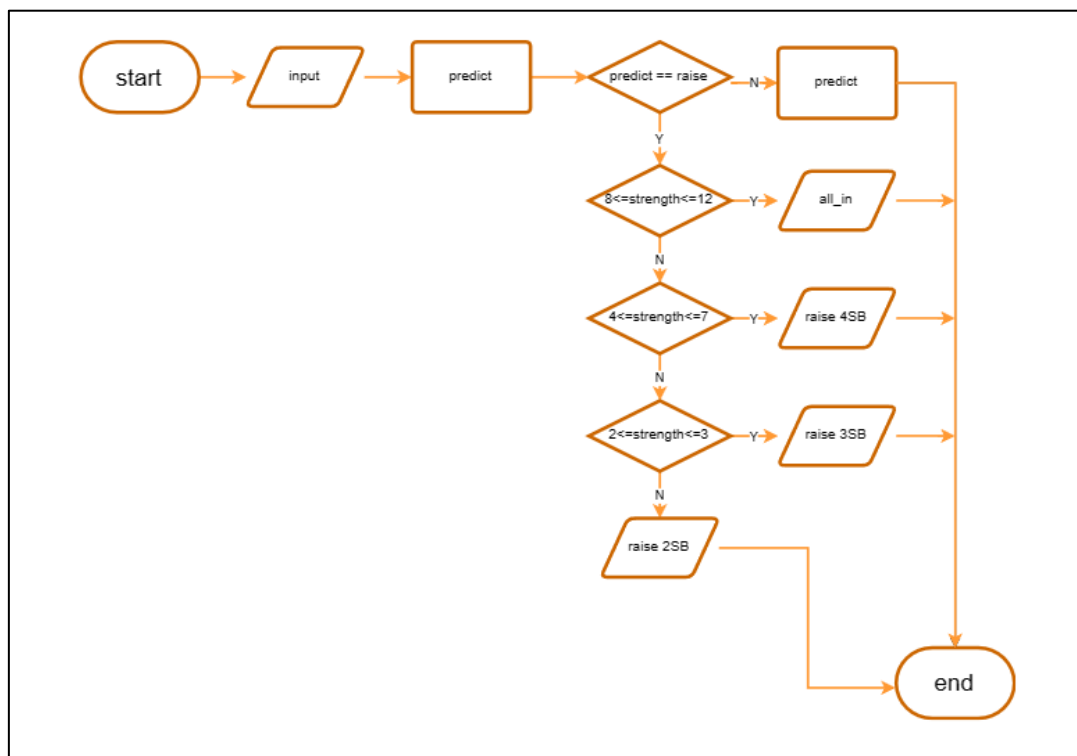


圖 7. All_in 與 Raise 流程圖

四、專題成果

4-1 真人 v.s AI

如圖 8 所示，在最新的 AI 對戰平台更新中，我們突破了傳統的一對一牌局限制，引入了更為複雜的多人牌局設定，使得遊戲更接近真實世界的多人撲克牌局環境，同時也提供了更多樣化的對戰選擇，玩家在遊戲開始前可自由根據自己的偏好選擇一對一到一對三的牌局形式，這樣的自由度不僅增加了遊戲的可玩性，也讓玩家有機會練習和提高自己在不同對局中的遊戲策略和技巧。我們還加入了新的 AI 模型，這些模型經過了從第 111 屆至第 113 屆的數據訓練，這些 AI 在遊戲策略上的多樣性和適應性，能夠更好的模擬真實玩家的行為，從而提供更加豐富和挑戰性的遊戲體驗。此外，這些新功能的加入同時也為我們提供了更多的數據，有助於我們分析和理解玩家行為，以進一步優化 AI 模型的表現，也推動了 AI 在遊戲領域的發展。通過這些創新，我們期待創造一個更加智能和互動的遊戲環境，為玩家帶來更加真實和刺激的牌局對戰。



圖 8. 真人對 AI 介面示意圖

4-2 AI v.s AI

隨著 AI 的對戰平台進入全自動化的新階段，我們將半自動牌局的設定更新，使其能夠完全自主運作，進行連串的對局並產生實用的數據，這不僅減少了繁瑣的人工操作，也讓數據收集變得更加迅速和精確。如圖 9 所示，我們新增的介面允許使用者從近三屆的 AI 模型中選擇，並可依據需求設定遊戲參數。每一輪的對戰都能在結束時展示每位 AI 玩

家的籌碼變化，這對於分析 AI 的行為模式十分有幫助，不只增進了遊戲的多樣性、提升了對局的效率，也為我們提供了一個更佳的實驗場域，用以測試和改進 AI 的決策能力。

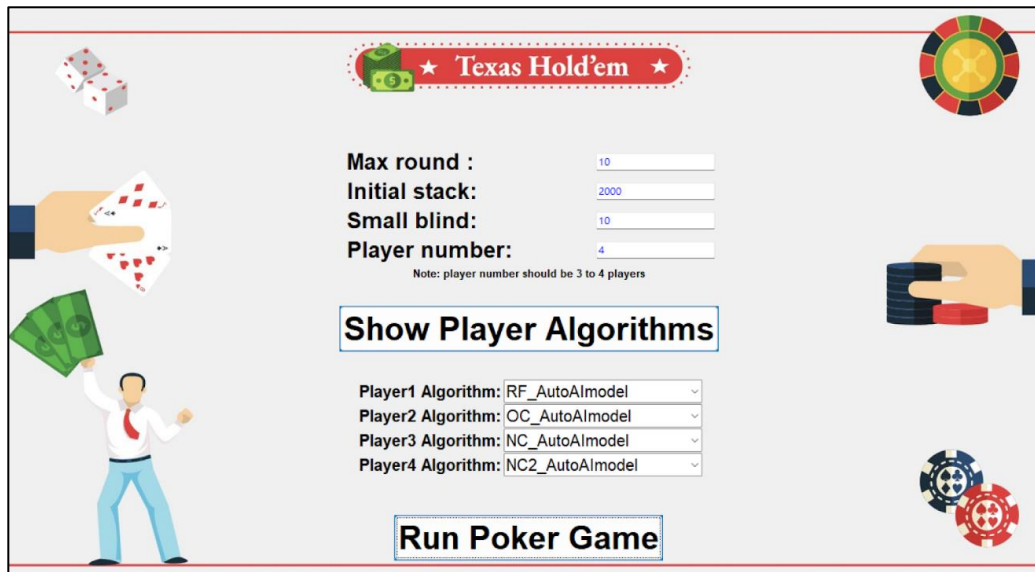


圖 9. AI 對 AI 介面示意圖

4-3 自動生成訓練資料

在第 113 屆 AI 模型 NewCNN 的設計中，我們著重於實務應用，讓系統能夠自動記錄三人或四人牌局的數據。這個模型採用了一種直觀的數據記錄格式，讓每一筆牌局資訊都能以二維陣列的形式被攤平並儲存，這使得後續的數據處理變得更為簡便和高效。這種數據記錄方式不僅簡化了傳統複雜的數據整理工作，還確保了數據結果的多樣性和豐富性。我們的目標是讓數據收集不再是研究人員和遊戲開發者的負擔，而是一個可以輕易存取、分析和應用的資源。透過這種自動化的數據收集過程，NewCNN 模型能夠更加精準地學習牌局中的各種情況，從玩家的策略選擇到下注模式，每一細節都能被模型學習和模擬。這對於模型的未來訓練和優化具有重要意義，因為它為 AI 提供了一個更加真實的學習環境，使其能在遊戲中作出更接近人類玩家的反應和決策。此外，這種自動化的數據記錄策略還為 AI 對戰平台帶來了新的可能性。隨著數據量的增加，AI 模型能夠接觸到更多樣化的遊戲情境，這對於提升 AI 的遊戲水平至關重要。同時，這也為我們提供了一個豐富的數據庫，供未來各種研究和分析使用，從而深化我們對於德州撲克這類策略遊戲的理解。

4-4 數據結果

• 玩家籌碼變化分析

在圖 10 (a) 所示的十輪牌局中，不同 AI 模型的籌碼變化呈現出顯著差異。RF 模型在初始階段快速累積籌碼，但第五輪後急遽減少，透露出可能過度激進且風險管理不足的策略。雖然此策略在開局可能帶來高收益，但長期看來，其表現缺乏穩定性，易於一次失誤導致大量損失。相反，OC 模型顯示出一個穩健增長的趨勢，但後期未呈現出顯著的領先優勢。對 OC 而言，過於保守可能未能充分利用手中牌的優勢，錯過增加籌碼的機會。與此相比，NC 及 NC2 模型展現出風險與收益間更平衡的表現，特別是 NC2 在牌局後期籌碼顯著增加，這反映了其策略的持續性和高效性。

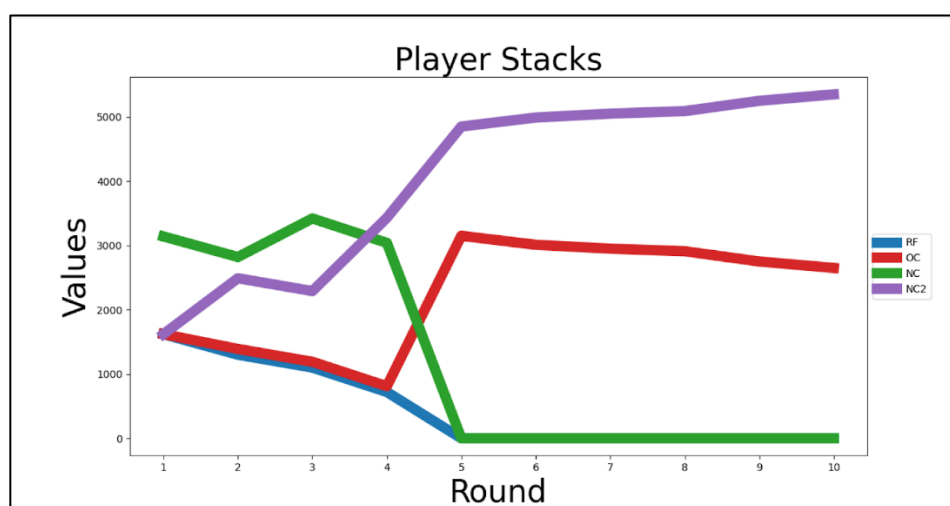


圖 10 (a). AI 對 AI 數據結果

• 遊戲互動性與策略調整

從圖 10 (b) 的數據波動中可見，牌局的互動性及各 AI 模型的適應性。RF 模型展現出最大的波動性，暗示其在遭遇不同對手策略時可能的不穩定性。而 OC 模型表現出一致性不足，顯示需要對策略進行改進以適應多變的遊戲情境。NC 及 NC2 的波動較小，顯示出在不同遊戲條件下的策略一致性，以及在面對多樣化對手時保持籌碼水平的能力。

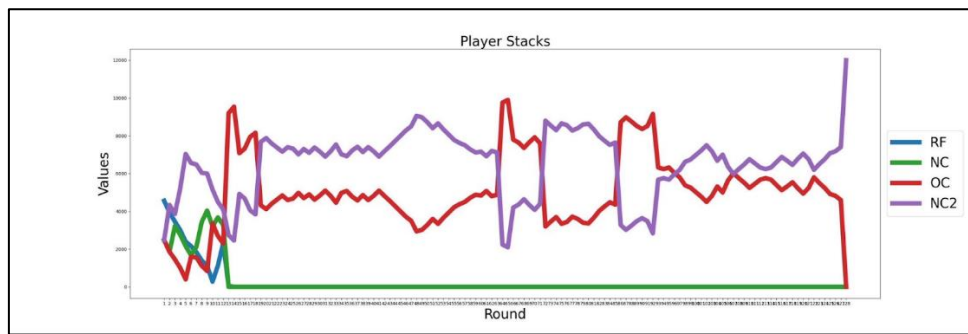


圖 10 (b). AI 對 AI 數據結果

• **策略多樣性與應對複雜性**

圖 10 (c) 展現了在較長牌局周期中的 AI 模型表現。NC 模型在開局階段領先，但後期被 NC2 超越，顯示 NC 模型可能採取了積極攻擊策略。NC2 的後期反超顯示了其策略的耐心與有效性，尤其在面對不確定性和長期競爭時。改進策略的複雜性方面，NC 模型需要改善長期策略以避免後期被超越。NC2 則需在早期階段強化攻勢，以確保在整個牌局中的競爭力。

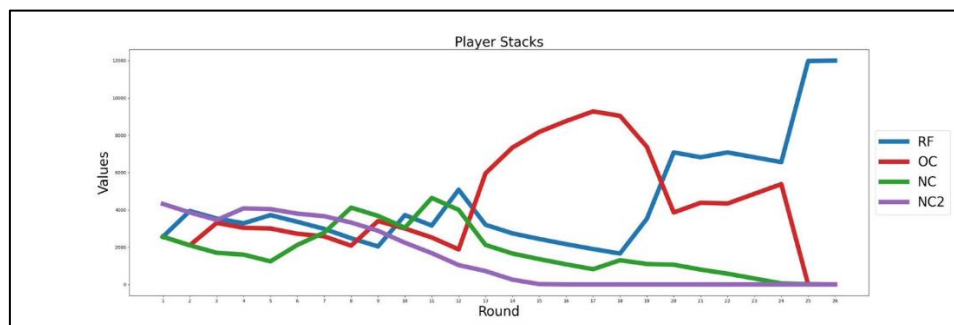


圖 10 (c). AI 對 AI 數據結果

五、 未來展望

5-1 往強化式學習導向發展

利用本專題現有之資料與成效，從遊戲環境中探索可能性，並考慮遊戲中玩家們與 AI 的預測行為與實際行為，來從中研究最佳化策略。將訓練時期拉長，以長期訓練之反饋來規劃遊戲走向，不限於勝率提升，利用強化式學習兼具探索與利用之特性，開創新穎之遊戲格局。

5-2 使用新資料集，蒐集更完整的數據庫

除了既有的資料集，未來預計採用更豐富的資料集，以獲得更多的遊戲數據、玩家行為、策略等信息，以建立更全面、多樣化的資料庫。這將為未來的研究和模型訓練提供更多的素材和可能性，有助於提高模型的泛化能力和預測準確度。

5-3 紀錄 AI 與 AI 對局數據 生成新的資料集

目前已成功生成 6 萬筆 AI 對 AI 多人德州撲克牌局的數據資料，可於未來進一步分析 AI 在德州撲克中的行為模式、策略選擇以及決策過程，並對 AI 模型進行更精確的調整與優化。

5-4 開發一個以上真人玩家對局模式

開發一個允許多個真人玩家參與的對局模式，這樣的模式可以提供玩家更多互動和挑戰，同時也能讓 AI 在與真人玩家對戰中獲得更多學習與提升的機會。

5-5 進一步優化 AI 模型的風險管理

在面對高不確定性的牌局中。增強模型對於多樣化對手策略和遊戲進程變化的適應性及彈性。引入更豐富的實際遊戲數據於 AI 訓練中，以提升模型在真實遊戲環境下的表現。

參 考 文 獻

- [1] Fault Diagnosis of Rolling Bearing Based on Tunable Q-Factor Wavelet Transform and Convolutional Neural Network [\(PDF\) Fault Diagnosis of Rolling Bearing Based on Tunable Q-Factor Wavelet Transform and Convolutional Neural Network \(researchgate.net\)](#)
- [2] [實戰系列] 使用 Keras 搭建一個 CNN 魔法陣（模型）[\[實戰系列\] 使用 Keras 搭建一個 CNN 魔法陣（模型） - iT 邦幫忙::一起幫忙解決難題，拯救 IT 人的一天 \(ithome.com.tw\)](#)
- [3] 決策樹 (Decision tree) [\[Day 12\] 決策樹 \(Decision tree\) - iT 邦幫忙::一起幫忙解決難題，拯救 IT 人的一天 \(ithome.com.tw\)](#)
- [4] ML 入門（十七）隨機森林(Random Forest) [ML 入門（十七）隨機森林\(Random Forest\). 介紹 | by Chung-Yi | 程式設計之旅 | Medium](#)
- [5] 【Day 15】學習率 Learning Rate [【Day 15】學習率 Learning Rate - iT 邦幫忙::一起幫忙解決難題，拯救 IT 人的一天 \(ithome.com.tw\)](#)