



# 陳鄚鄚

DAN-YI, CHEN

跨領域 跨國際 跨產學

## 教育經歷

國立陽明交通大學 (原交通大學)  
生物科技學系 2020 - 至今

計算生物與生物資訊專業  
校級人工智慧跨域學程 - 生醫組

德國柏林工業大學  
生物科技學系 2023 - 2024

## 在校成績

- GPA (3.76 / 4.3)
- 系排 (17 / 73) (23.29%)

## 專長技能

- 生物資訊學
- 機器學習
- 生醫大數據
- 生醫訊號處理

## 程式語言

- Python (良好)
- C++ (中等)

## 外語能力

- 英文 IELTS 6 (良好)
- 德文 B2 (良好)

## 實習經歷

工業技術研究院

2024 - 至今

資訊與通訊研究所

國科會晶創臺灣方案:

應用於腦神經退化疾病之晶片共通平台開發

• 企劃書撰寫:

技術說明、競品及專利分析。

• 即時數據處理與分析:

利用 Socket 實現嵌入式系統和用戶端之間的即時數據傳輸，並應用 threading 模組處理手動或自動狀態下的數據輸出。

• 機器學習模型構建:

精確預測帕金森氏症患者施予深腦電刺激治療的時機點，測試集的準確率高達 92.7%。

• 數據可視化:

利用 Plotly、Matplotlib 開發即時數據的圖表可視化工具，即時顯示腦波及 PSD 訊號。

• 用戶端界面與前端設計:

使用 Pywebview、JavaScript 和 HTML 整合前後端數據並即時顯示多種圖表。

## 專題研究

智慧型計算實驗室 - 何信瑩教授

研究題目：一種新型與轉移相關的 lncRNA 特徵組，用於預測皮膚黑色素瘤的預後

• 特徵篩選與基因算法應用:

利用秩和檢驗和基因演算法選出由轉移相關的 21個 lncRNA 組成的最佳特徵組。

• 設計風險評分系統:

基於最佳特徵組設計的風險評分系統，精確地將患者分成不同風險組。

• 預後模型開發及驗證:

開發 CoxPH 與 SVM 模型預測不同風險組患者的存活及轉移風險。

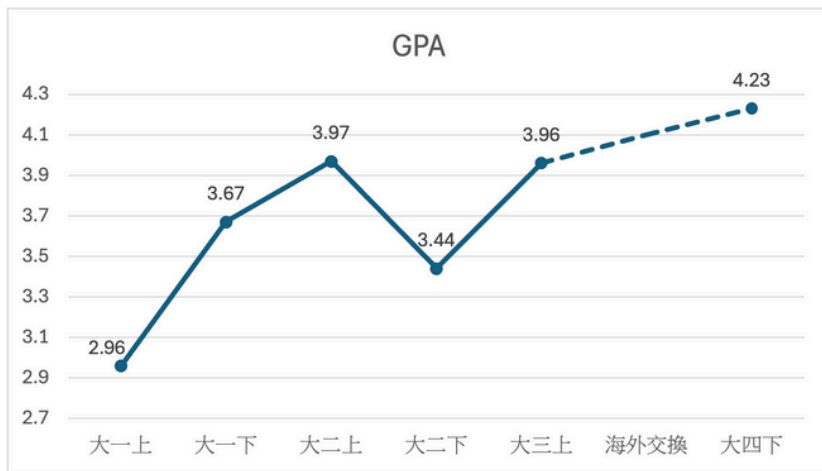
• 生物學意義分析:

用特徵敲除分析鑑別出對轉移預測貢獻最大的 10個 lncRNA，提出潛在的生物標記物。

# 個人背景

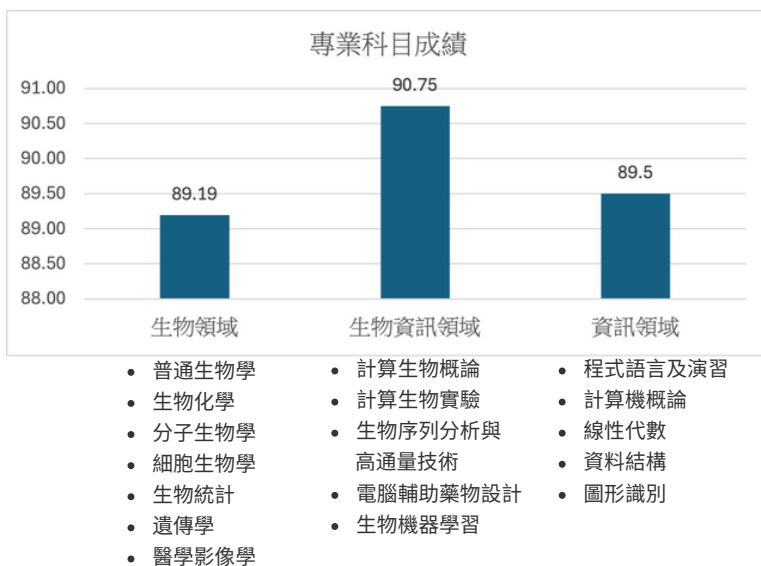
我在一個溫暖的基督徒家庭中長大，從小父母就鼓勵我**獨立思考、勇於跳脫舒適圈**嘗試新事物，藉此培養了我的好奇心以及勇於面對挑戰的個性，使我總是**帶著熱情與衝勁發展多元的興趣和專長**。

## 學習歷程



從大一入學起，我就專注於修習系上**計算生物與生物資訊的專業課程**，並申請**校級人工智慧跨域學程-生醫組**，希望能將 AI 和大數據的技術靈活應用於生醫領域，成為資訊與生醫領域之間的橋樑。

一開始大一的基礎課程對我而言很具有挑戰性，但隨著漸漸熟悉大學上課的步調及讀書方法，我在生物及資訊領域的成績也有顯著的提升。直到大二的下學期，因著修習四門生物及資訊的核心課程，並同時準備德文和英文檢定考試，以預備交換學生的申請，才無法兼顧課業及檢定，使得成績下滑。



我在資訊及生物領域都有很平衡的學習與表現，平均都獲得**接近90分**的成績，也修習了各個領域的實驗或實作課程，從而累積實務經驗。

# 學習歷程

我利用三年寒暑假的時間密集地學習德文，通過**歌德 B2 德文檢定**，並獲得到德國柏林工業大學交換一年的機會。在這一年裡，提供我良好的語言學習環境，也開拓了我的視野，使我學習用不同文化的角度來思考。此外，我也不斷挑戰自己，嘗試接觸陌生的環境以及學習新技能，培養了我在困境中尋找解答的毅力，以及**危機處理**的能力。這樣**快速學習**以及**靈活的適應能力**，幫助我融合不同領域的概念與技術，也能**迅速掌握不熟悉的事物**，使我更適合投入生物資訊領域的研究中。

為了加強資訊領域的能力，我在柏林工業大學交換一年中修習了許多軟體工程與理論電腦科學的課程，不僅加強了我的專業能力，更培養我**自主學習**的能力以及**積極的學習態度**，使我在回國後獲得**班排第一**的佳績。



於柏林工業大學與同學合影

在專題研究中，我獨立完成了研究的各個環節，涵蓋題目設計、數據收集及分析、模型開發等，使我打下扎實的基礎，並培養**獨立進行完整研究**的能力。我的指導教授對於研究的標準很高，訓練我**高效地管理研究進度**，並藉由大量閱讀相關的論文來學習其他學者的成功經驗，從而優化我的研究設計。這個經驗不僅讓我對機器學習及特徵篩選有深入的理解，也讓我學習在實驗室的討論中保持批判性的思維，找到突破現有研究的創新點。

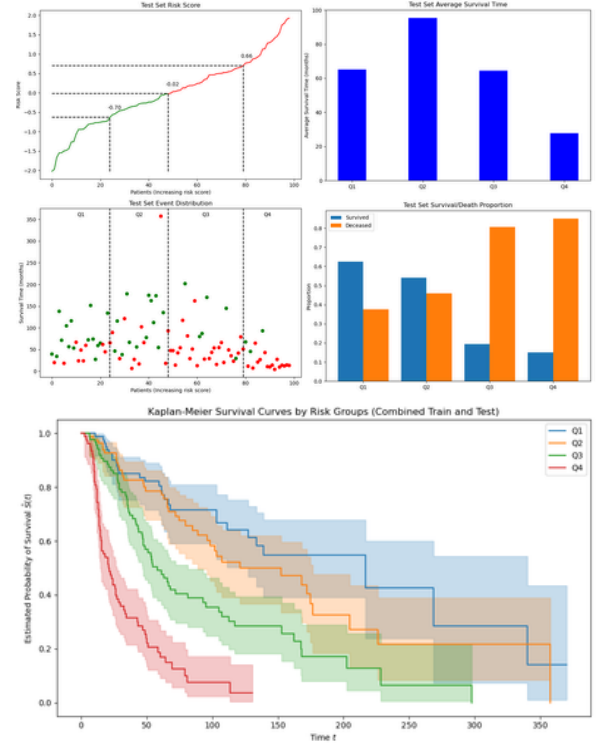
我希望能夠累積更多在資訊領域的實務經驗，所以在今年暑假申請工研院資通所的實習機會。我主要負責研究**國科會的晶創臺灣方案**的模型開發，過程中我表現出**高效的學習力及適應力**，只用一個多月就完成今年度的專案目標，並獲得主管及醫師的肯定。此外，我也同時協助院內多項前瞻計畫，涵蓋**生物、資訊、醫學及金融**等領域，訓練我具備跨領域的思維與分析能力。這段經歷讓我學習有效率地管理不同專案，不僅加強了我的研究及溝通能力，也更了解業界對於產品及人才的要求。

**我相信我具備跨領域、跨國際、跨產學的經驗與思維，並富有熱情與毅力，能夠提供貴所研究上的突破與貢獻。**

## 專題研究 一種新型與轉移相關的 lncRNA 特徵組，用於預測皮膚黑色素瘤的預後

皮膚黑色素瘤(CM)是台灣十大癌症之一，其惡性高且易發生早期轉移，所以我提出了一個基於 lncRNA 的預後模型，用於預測 CM 患者的存活及轉移風險。我從 TCGA 數據庫獲得 CM 患者的 lncRNA 表達譜及臨床數據，先利用 log2 差異分析及變異數分析對 lncRNA 表達譜進行前處理，再利用了秩和檢驗篩及基因演算法 IBCGA 進行特徵篩選，最終選出由21個轉移相關 lncRNA 的最佳特徵組。

我基於最佳特徵組設計一種風險評分系統，以精準地將患者分成不同風險組，再利用 Cox-CM 模型和 SVM-CM 模型預測患者的存活及轉移風險，**實現個人化的預後診斷**。Kaplan-Meier 生存曲線顯示高風險組患者的存活率顯著低於低風險組 ( $p < 0.0001$ )，**證實風險評分在預測存活時間上的可行性**。最終我透過特徵敲除分析鑑定出對轉移預測貢獻最大的10個 lncRNA，其中有部分 lncRNA 已被證實與癌細胞的轉移有關。

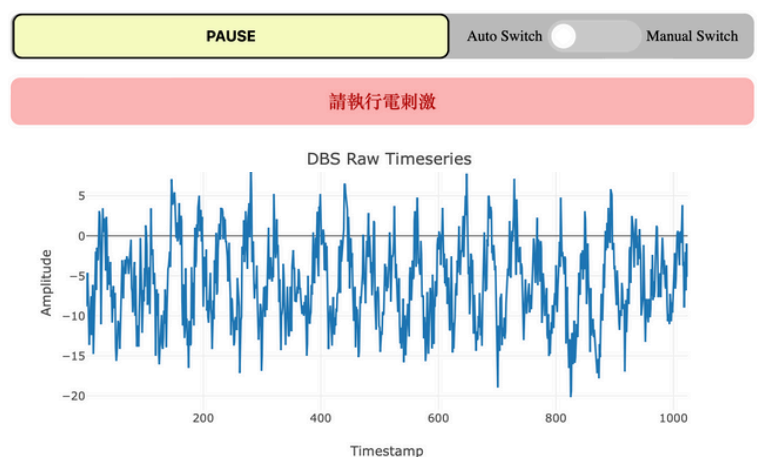


## 實習成果 應用於腦神經退化疾病之晶片共通平台開發

本研究的貢獻在於透過機器學習分析模型，**適時且精確預測帕金森氏症患者施予深腦電刺激 (DBS) 治療的時機點**，以減少患者的副作用、治療成本及手術風險。我從 MRC Brain Network Dynamic Unit 下載了26個病患進行 130 Hz DBS 前後的 STN 局部場電位數據，先利用快速傅立葉變換、功率譜密度計算進行前處理，再利用了秩和檢驗提取具有統計顯著性的頻譜特徵 ( $p < 0.005$ )，用來訓練 XGBoost 模型，最後將訓練好的模型移植到嵌入式系統，並即時顯示腦波訊號、頻段功率及能量分布。

本模型在測試集的**準確率高達 92.7%**，AUC 為0.971，**優於現有的研究成果**(準確率87%)。期望能突破現有技術的不足，為帕金森氏症患者提供更安全、有效的治療方法。

DBS Data Streaming Demo



## 申請動機

### 為什麼選擇陽明交大數據科學與工程研究所?

我對資訊與生醫應用的結合充滿興趣與熱情，並有多元的研究和實務經驗，我希望能申請陽明交大數據所，在數據工程領域進一步深耕相關的理論知識與應用技巧，用來處理生醫領域大量且複雜的資料，為社會的生醫領域做出貢獻。

#### 1. 提升資訊領域的能力

在專題研究及實習的過程中，我深刻體會到人工智慧及大型語言模型在生醫領域應用上的潛力，然而我發現在處理大量數據和資料時，我對於相關的理論及實作經驗是不夠的，貴所所能夠提供我最頂尖的師資及教學資源，提供資料探勘或演算法相關的進階課程，並應用於生醫領域，藉此拓展並加深我的專業能力。

#### 2. 整合跨領域專業

貴所師資的研究領域結合了生物資訊、資料探勘與巨量資料分析等多元領域，如胡毓志教授將機器學習預測模型應用於免疫、疾病和藥物等領域，藉由分析免疫反應實現疾病的早期診斷，或是預測藥物的使用劑量或副作用等，我希望學習教授的研究技巧並多方累積實務經驗，培育我將資訊領域創新的技術應用於疾病或癌症治療，在個人化精準醫療的領域發展創新的技術與突破。

#### 3. 陽明交大卓越的國際聲譽

我在交大經歷大學階段的學習與專題研究中，受到學校對於理論與實作並重的栽培，使我在德國交換及實習時獲得許多肯定與信賴。我希望能在貴所繼續精進自己，培養生醫領域的知識和技能，為未來的職涯發展打下穩固的基礎，並在求職或出國進修時具備足夠的競爭力。

## 申請動機

### 為什麼我適合就讀陽明交大數據科學與工程研究所?

#### 1. 符合跨領域背景與研究經驗

我具備計算生物與生物資訊的專業，並完成人工智慧跨域學程，使我擁有結合資訊與生醫領域的能力。此外，不管是在皮膚黑色素瘤預後模型的專題研究，或是在實習期間閉迴路深腦刺激的系統開發，都是利用生物資訊的技術進行數據分析與模型開發，這高度契合貴所對於跨領域應用的需求，以培養創新研發的能力。

#### 2. 軟體知識與數據分析能力

在求學以及海外交換期間，我修習了許多軟體工程與數據分析的相關課程，本學期也額外加強了演算法及作業系統相關的課程，並在專題與實習中靈活應用，能夠結合所學的專業技能並規劃專案。

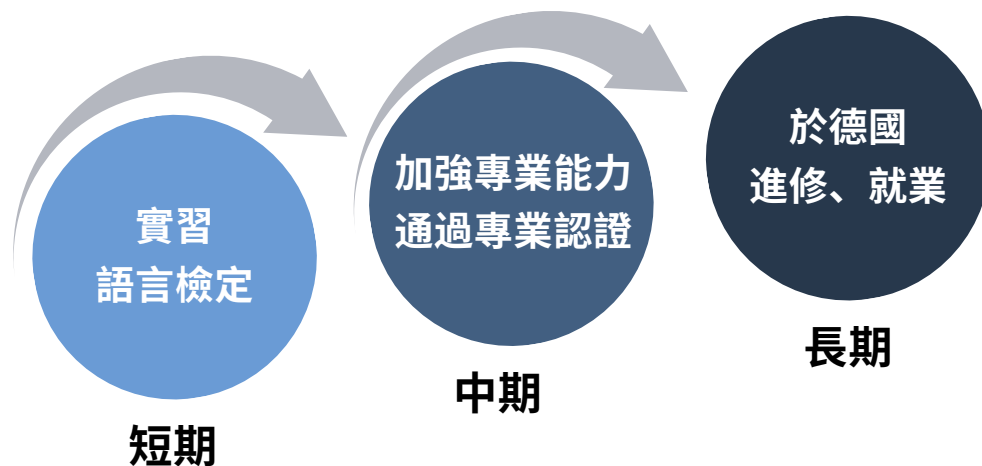
#### 3. 獨立研究與創新能力

在專題研究和實習中，我具備獨立設計題目、分析並完成研究目標的經驗，在過程中培養了創新思維和解決問題的能力，並努力呈現最優化的結果。這樣的獨立研究經驗有助於我在貴所的學習中不斷創新，面對研究中的各種挑戰。

#### 4. 多元語言優勢

我在大學期間特別重視德文和英文的能力培養，並具備在歐洲生活一年的經驗。在專題研究及實習中也培養閱讀論文的習慣，不僅有助於閱讀及撰寫論文的效率，也利於未來出國參與研討會或進修。

## 學習計畫



### 短期 (1年內)

#### 1. 工研院實習：

- 完成專案今年度的計畫，也能嘗試實作其他領域的專案進度，從中深入學習生醫領域的數據分析及程式開發，增加實務經驗。

#### 2. 預習研究所課程：

- 加強演算法以及計算機相關的理論及實作，以銜接貴所的進階專業課程。

#### 3. 德文與英文檢定考試：

- 德文 C1 考試：準備 C1 等級的德語檢定，達成申請德國研究所的語言門檻。
- 英文檢定：準備 IELTS 檢定，目標考取7分以上的成績。

#### 4. 完成國家義務役：

- 並利用空閒時間進行學科或語言的準備，確保順利銜接未來的學業規劃。

## 學習計畫

### 中期 (就讀碩士期間)

#### 1. 完成碩士研究：

- 將機器學習或大型語言模型應用於基因體學、癌症生物學等，實現個人化精準醫療。

#### 2. 加強專業能力：

- 資訊領域：如深度學習、嵌入式即時系統，以及演算法相關等進階或核心課程。
- 生醫領域：如生醫資料探勘、生醫資訊深度學習等課程。

#### 3. 參加全國大專院校人工智慧競賽或企業黑客松：

- 積極參與各類人工智慧競賽或黑客松活動，實際應用所學的技術，並藉此累積更多 AI 和生物資訊相關的競賽經歷。

#### 4. 擔任助教或研究助理：

- 透過指導他人來實際消化所學的專業知識，並發掘自身的學習盲點。也能透過協助教授研究的過程中，提升自身的研究能力及溝通技巧。

#### 5. LeetCode 練習：

- 系統性地練習 LeetCode 中的演算法與資料結構題目，不斷提升程式設計和解題能力，為未來考取專業認證及求職做好充分的準備。

### 長期 (碩士畢業後)

#### 1. 申請德國碩士或博士學位：

- 德國慕尼黑工業大學在資訊科學領域提供豐富的教學資源及跨國合作機會，我也計畫至生資領域的企業實習，並在國際期刊發表論文。

#### 2. 程式語言與專業認證：

- 考取 ITS Python 核心能力認證，並通過 ITS 及 TIPCI AI 人工智慧能力認證。

#### 3. 在德國生資領域公司就業：

- 畢業後希望能在德國的生物資訊公司或研究機構就業，發揮跨領域整合的能力，也希望進一步回饋台灣的社會及生資產業的發展。

## 研究計畫

### 腦神經退化疾病的即時生理訊號 CNN-GRU 預測模型開發

#### 摘要

我希望提出了一種針對帕金森氏症 (PD) 患者，基於即時訊號預測病徵發生的創新模型。相比傳統的腦電圖 (EEG) 預測模型只針對  $\beta$  波爆發的檢測，此研究透過 CNN-GRU 模型融合了多種生理訊號，包括 EEG (腦電圖)、EMG (肌電圖) 和 ECG (心電圖)，以提高預測的準確性，並採用生成對抗網路 (GAN) 進行數據擴增，增加模型的泛用性。此外，該模型使用可解釋性 AI (XAI) 技術，以確保在臨床應用中的透明性及可解釋性。(完整內容請參考其他有利審查資料)

#### 研究方法

##### 1. 數據集收集與前處理

從 MRC Brain Network Dynamic Unit 和 PhysioBank 獲取 EEG、EMG 和 ECG 等生理訊號數據，並使用快速傅立葉變換、功率譜密度和標準化方法進行前處理，篩選出具有統計顯著性的特徵，並力用 GAN 來合成多模態生理數據，以解決數據難以取得的問題。

##### 2. 模型架構

此研究使用 CNN-GRU 混合神經網路，CNN 層負責提取空間特徵，而 GRU 層負責捕捉信號中的長短期時間依賴性，進而即時預測  $\beta$  波爆發和其他病徵。

##### 3. 可解釋性與臨床實用性

利用 XAI 技術的 Layer-wise Relevance Propagation (LRP) 方法，分析每個特徵對模型預測的貢獻程度，從而提高模型的透明度和可信度，增加在臨床應用中的可行性。

#### 創新性與貢獻

##### 1. 多模態生理訊號結合：

此研究結合 EEG、EMG 和 ECG 多模態訊號，能夠更準確地預測 PD 病徵，以提升準確率。

##### 2. CNN-GRU 混合模型：

透過 CNN-GRU 模型，能夠有效地提取生理訊號中的空間和時間性特徵，進而提高即時預測關鍵事件發生的準確率。

##### 3. GAN 數據擴增：

應用 GAN 技術解決臨床數據不足的問題，以提升模型的泛化能力。

##### 4. XAI 技術：

整合 LRP 方法，使模型預測結果具有可解釋性，增加在臨床應用中的可性度及實用性。