

主導課程四：生成式 AI：文字與圖像生成的原理與實務 (Generative AI: Text and Image Synthesis Principles and Practice)

課程基本資料

開設學校：政治大學

開授教師：蔡炎龍

班級人數：約2000人 (保留開課學校 500 人)

開課級別：碩士課程 (政大學碩合開)

授課語言：中文

授權方式：條件式授權

建議協同教師學經歷：具深度學習基本概念，對生成式 AI 熟悉、開過相關課程教師佳，但非必要。

同步遠距上課時間：每週二 16:00-19:00

遠距上課位置：Facebook【政大應數系直播中心】

<https://facebook.com/groups/nccumathonline/>

課程網頁：

修課人數與助教比例：每 25 名學生需 1 名助教 (建議)

課程概述

「生成式 AI：文字與圖像生成的原理與實務」是一門兼具理論深度與實作樂趣的課程，專為希望深入了解生成式 AI 的技術與應用的學生而設計。不論對 AI 的認識是基礎還是進階，我們都希望透過這門課程，帶領同學探索生成式 AI 的無限可能。學生將會學習神經網路、

GAN、Transformer、大型語言模型、RAG、AI Agents、Diffusion Models 等技術，並運用工具如 OpenAI API、LangChain、HuggingFace 及 AutoGen 等等，完成從文字生成到圖像生成的多樣應用。

課程目標

1. 理解生成式 AI 的核心技術，包括神經網路、GAN、Transformer、大型語言模型、RAG、AI Agents、Diffusion Models 等。
2. 實際運用各種工具和框架，例如 OpenAI API、LangChain、AutoGen、HuggingFace、Foocus，打造多樣的生成式 AI 應用。
3. 探討生成式 AI 的社會與倫理挑戰，從技術層面與實務層面提出創新解決方案。
4. 完成期末專題，整合所學內容，設計並展示一個實用的生成式 AI 系統。

課程特色

循序漸進的內容設計：從神經網路的基礎原理出發，逐步深入進階模型與應用。

實作為主，理論為輔：課堂中將使用 Colab 進程式實作，配合具挑戰性的課後作業，邊學邊做。

探索最新技術：涵蓋最新的生成式 AI 模型與工具，掌握 AI 發展趨勢。

多元化的應用場景：課程內容涉及文字生成、圖像生成、對話機器人、Agentic AI 等多個領域。

倫理與應用並重：不只在技術方面，更引導反思生成式 AI 的社會影響與，強調「負責任地使用 AI」。

參考書目

主要是上課講義，其餘參考資料於課程中介紹

課程內容大綱

週次	日期	課程內容	備註
1	2/18	課程介紹與生成式 AI 概述	課程目標與內容簡介，為什麼要研究生成式 AI，介紹 Colab 平台的基礎操作
2	2/25	神經網路的概念	簡介神經網路的核心概念（感知器、多層感知器）、激發函數與反向傳播，實作一個簡單的 MNIST 手寫數字分類
3	3/4	生成對抗網路 (GAN)	介紹曾經被當作生成式 AI 希望的生成對抗網路 (GAN)，包括 GAN 的生成原理、著名的應用範例，及為什麼暫時不再是主流的原因
4	3/11	文字生成 AI 與大型語言模型	文字生成 AI 的基本概念，包括 RNN 及 transformers 有「記憶」的神經網路簡介，及 seq2seq 模型、詞嵌入等等原理
5	3/18	RNN 及 transformers 的數學原理	RNN 及 transformers 的數學基礎與架構，複習基本矩陣運算、瞭解注意力模式的原理
6	3/25	大型語言模型 (LLM) 的應用及倫理議題的挑戰	著名的大型語言模型 (LLM)，常見應用，及倫理議題的討論
7	4/1	打造自己的對話機器人	使用 OpenAI API，用程式的方式打造自己的對話機器人
8	4/8	檢索增強生成 (RAG) 的原理及實作	RAG 的概念、資料檢索與生成的結合方法，實作基於 LangChain 的 RAG 系統
9	4/15	專家講座	業界專家分享 RAG 在金融上的應用與挑戰

10	4/22	Agentic AI 與 AI Agents	什麼是 AI Agents ? 設計目標導向的 Agent · 介紹 AutoGen 框架並展示應用案例
11	4/29	變分自編碼器 (VAE) 開始的冒險旅程	解釋 VAE 的原理 · 及為什麼這相對簡單的模型 · 會成為圖像生成 AI 的重要想法
12	5/6	Diffusion Models 與圖像生成	介紹 Diffusion Models 的原理 · 包括加噪過程 · U-Net 架構的減噪等概念
13	5/13	文字生圖 AI 的原理及實作	介紹「理解」使用者輸入文字的 CLIP 模型 · 著名的圖像生成 AI 模型及運用方式
14	5/20	Diffusion Models 進階主題	Latten Diffusion Models, LoRA, ControlNet 等概念
15	5/27	Foocus 實作圖像生成	介紹開源 WebUI 框架 Foocus · 使用 Stable Diffusion 等開源模型做圖形創作
16	6/3	研討會型式的期末專題成果分享	

成績評量方式

* 各校評量方式以各校老師公告為主

- 作業及反思: **75%**

每次作業繳交時間在兩週內，每次作業滿分為 **10** 分。歡迎同學運用大型語言模型協助，但只能寫出下一次 **prompt** 就生得出結果水準的作業，得分最高 **3** 分。

- 期末專案: **20%**

每個人需完成一個生成式 **AI** 應用專案。期末分享採研討會模式，以投稿方式參與，獲選同學參加期末專案分享，並有額外加分。

- 上課參與 **5%**

非常鼓勵同學在「直播」時間參與課程，另外上課互動、參與討論等等，皆會列入考量。

- 額外加分

課程有「閃電秀」安排，同學們可自由報名，在每次上課第三節時，以 5 分鐘內的時間，分享自己對生成式 AI 相關的心得。每次上課最多可接受 5 位同學的分享。

課程要求

1. 會使用 **Google Colab** 雲端運算平台，請同學準備好自己的 **Google** 帳號。本課程的作業應該免費版就足夠，但可以考慮自己狀況是否升級。
2. 建議 (非要求) 於 <https://platform.openai.com/> 儲值使用 **OpenAI API** 的 **credit**，應該 5 美金就完全足夠課程的需求。課程中我們還是會提供其他免費的方案，但使用 **OpenAI API** 可能會比較方便 (特別對技術不是那麼熟悉的同學)。
3. 非常強調不可以抄襲，包括抄襲網路上的作品，或者直接抄襲生成式 **AI** 產出者，皆是不可接受的。本課程是生成式 **AI** 課程，使用大型語言模型協作，不但是允許，甚至是鼓勵的。這裡的抄襲是直接下一個 **prompt** 就能產出的結果，直接當作業是不能接受的。

課程基本資料

開設學校：國立清華大學

開授教師：胡敏君

班級人數：約1000人 (保留開課學校 60 人)

開課級別：學碩合開

授課語言：中文

授權方式：條件式授權

建議協同教師學經歷：資工、電機、動機相關科系畢業。

同步遠距上課時間：每周一晚上6:30~9:20

遠距上課位置：國立清華大學台達館106教室（遠距連結 meet.google.com/wbh-oihg-jsn）

課程網頁：

修課人數與助教比例：每 10 名學生需 1 名助教

課程概述

本課程模組分為三個主要的部分，分別為即時追蹤與地圖建置(**SLAM**)、基於機器學習之場景理解(**Scene Understanding**)與探索導航的動作控制(**Action Control**)。即時追蹤與地圖建置部分包含機率模型與相機模型等理論基礎，也包含基於深度學習之**RGB-based**的**3DSLAM**方法。場景理解的部分包含機器學習的基本概念，再帶到深度學習的技術與目前的物件偵測與語意切割技術。動作控制的部分則包含路徑規劃與導航演算法，並帶入強化學習的概念來引導行進的路徑。

參考書目

- Richard S. Sutton and Andrew G. Barto, Reinforcement Learning: An Introduction, Second Edition, MIT Press, Cambridge, MA, 2018
- Sebastian Thrun, Wolfram Burgard, and Dieter Fox, Probabilistic Robotics, 2005. (Intelligent Robotics and Autonomous Agents series)
- Kevin Murphy, Machine Learning: A Probabilistic Perspective.
- Daphne Koller and Nir Friedman, Probabilistic Graphical Models: Principles and Techniques, 1st Edition, 2009.
- Ian Goodfellow, Yoshua Bengio and Aaron Courville: Deep Learning.

課程內容大綱

週次	日期	課程內容	備註
1	2/17	Introduction to Robotic Navigation and Exploration	
2	2/24	Kinematic Model and Path Tracking Control * Control System Basics * PID Control * Basic Kinematic Model * Differential Drive Vehicle * Pure Pursuit Control * Kinematic Bicycle Model	Lab 1
3	3/3	Motion Planning * Motion Planning Introduction * Path Planning * Curve Interpolation * Trajectory Planning	Lab 2

		* Path Planning	
4	3/10	Reinforcement Learning (I) * MDP * Value Function * Bellman Equation * Reinforcement Learning	
5	3/17	Reinforcement Learning (II) * Q-Learning / Sarsa / DQN * Policy Gradient / Actor-Critic	
6	3/24	Project Environment Building (I)	Lab3
7	3/31	Project Environment Building (I)	Lab4
8	4/7	Project Environment Building (III)	Lab5
9	4/14	SLAM Back-end (I) * State Estimation and SLAM Problem * Probability Theory and Bayes Filter * Kalman Filter / Extended Kalman Filter	
10	4/21	SLAM Back-end (II) * Graph based Optimization * Graph Optimization for 2D SLAM (Bundle Adjustment)	

11	4/28	3D SLAM (I) * Feature Descriptor * Multi-view Geometry * Lie Group & Lie Algebra	
12	5/5	3D SLAM (II) * 3D SLAM: ORB-SLAM * Direct Method * DNN-based SLAM	
13	5/12	3D Embodied Agent	
14	5/19	Paper Presentation (I)	
15	5/26	Paper Presentation (II)	
16	6/2	Project Presentation & Demo	

成績評量方式

- 作業: **60% (15% for each HW)**
- 論文閱讀報告(10%)
- 自走車期末專題(含實作、書面報告、口頭報告): **30%**

課程要求

- 建議學生需已修過Python程式設計、影像處理、深度學習。
- 學生須自備具GPU顯卡之電腦。
- 本課程期末專題採分組開發，為避免影響同組修課同學之權益，本課程不接受期中退選，請謹慎評估可投入的時間再選課。