GenAI HW2
都是AI的作文比賽

TA: 陳瑋智、謝承修、呂俐君

ntu-gen-ai-2024-spring-ta@googlegroups.com

Deadline: 2024/03/14 23:59:59 (UTC+8)

Outline

- Task Description
- Workflow
- Grading Policy
- Important Dates
- Regulations

Links

- [NTU Cool HW2 Forum](#)
- [MediaTek DaVinci](#)
- [OpenAI ChatGPT](#)
- [Anthropic Claude](#)
- [Google Gemini](#)
- [Microsoft Copilot](#)

Task Description

都是AI的作文比賽: 參賽者跟評審都是AI

There are two parts in this assignment: Writing essays and judging the essays

- You will use AI models to **generate** two essays
- You will use AI models to **evaluate** the essays

Workflow

Workflow

| 1. Generate an essay |

- Ask Language Models to write it
- Write it on your own
- Any other methods

Platforms
  a. OpenAI - ChatGPT
  b. MediaTek DaVinci（達哥）
  c. Others

| 2. Evaluate your Essay |

We use generative AI as a TA to evaluate your essay automatically.

Generative AI TA [批改助教] Info：
1. Platform: MediaTek DaVinci
2. Model: GPT-4
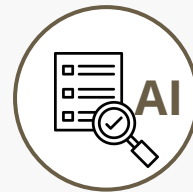
| 3. Submit to NTU COOL |

Part 1. Generate Two Essays with AI Models

Workflow

1. Generate an essay

- Ask Language Models to write it
- Write it on your own
- Any other methods

Platforms
- a. OpenAI - ChatGPT
- b. MediaTek DaVinci（達哥）
- c. Others

2. Evaluate your Essay

We use generative AI as a TA to evaluate your essay automatically.

Generative AI TA [批改助教] Info：
1. Platform: MediaTek DaVinci
2. Model: GPT-4

3. Submit to NTU COOL
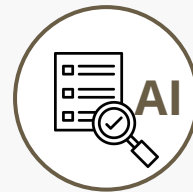
I. Short Essay Writing – English

Do you agree or disagree with the following statement?

Artificial Intelligence will eventually replace humans in most areas of work in the future world

Write an English essay with around 300 words to express your stance with specific reasons and examples to support your answer

II. Short Essay Writing – Mandarin

參考 113學年度學科能力測驗試題 國語文寫作能力測驗 第二部分 所附的文章,

請以「縫隙的聯想」為題, 撰寫一篇約五百字的繁體中文文章,

並結合生活經驗或見聞, 書寫你的感思與體悟。

Sample Methods to Prompt AI

- **Directly use the task description:**

> Write an essay with the topic {Topic}.

- **Describe your needs:**

> Write a **3-paragraph** essay about {Topic}. Please use a **professional tone**.

- **Role-playing:**

> You are a perfect writer. Please write an essay about {Topic}.

- **Any other creative prompts are welcome.**

<u>Replace {Topic} above with your chosen topic</u>

Short Essay Writing

- We expect you use generative models to complete this assignment,

- although it is permissible if you prefer to compose them manually

  - Editing the generated essays is allowed
  - Editing the essays that you wrote with AI is allowed

Part 1-a. Generate An Essay with ChatGPT

New chat

Click here to start a fresh conversation

ChatGPT 3.5 ⌄

⬡

How can I help you today?

Plan an itinerary
for a literary tour of England, visiting famous authors' ...

Design a database schema
for an online merch store

Show me a code snippet
of a website's sticky header

Create a charter
to start a film club

Message ChatGPT... **Type a message here to interact with ChatGPT**

Upgrade plan
Get GPT-4, DALL·E, and more

GE NewAccount

ChatGPT can make mistakes. Consider checking important information.

?

New chat

Today

Joable: Atom Trust Issues ···

ChatGPT 3.5 ⌄

You
Tell me a joke

✎ ← **Edit the message and regenerate a response**

ChatGPT
Sure, here's a joke for you:

Why don't scientists trust atoms?

Because they make up everything!

Copy the response → 📋 👎 ↺ ← **Regenerate ChatGPT's response**

Upgrade plan
Get GPT-4, DALL·E, and more

GE NewAccount

Message ChatGPT...

ChatGPT can make mistakes. Consider checking important information.

New chat

Today

Joke: Atom Trust Issues

Share link to Chat

Messages you send after creating your link won't be shared. Anyone with the URL will be able to view the shared chat.

You
Tell me a joke

ChatGPT
Sure, here's a joke for you:

Why don't scientists trust atoms?

Because they make up everything!

Joke: Atom Trust Issues ✎

January 27, 2024

More Info ⬈

Click here to copy →   🔗 Copy Link

Upgrade plan
Get GPT-4, DALL·E, and more

NewAccount

Message ChatGPT...

ChatGPT can make mistakes. Consider checking important information.

16

Part 1-a. Generate An Essay with ChatGPT: Demo

OpenAI ChatGPT Tutorial

Click here to load ChatGPT

ChatGPT ⬤

Suggest fun activities
for a family of 4 to do indoors on a rainy day

Get started

| Log in | Sign up |

Here to create your account

⬡ OpenAI

Terms of use | Privacy policy

19

Create your account

註冊 OpenAI 以繼續使用 Apps Client。

電子郵件地址
Enter your email here

繼續

已有帳戶？ 登入

─────────── 或 ───────────

G 繼續 Google

◼ 繼續 Microsoft Account

 繼續 Apple

or log in with existing accounts →

Create your account

為OpenAI設定密碼以繼續到

CreateNewAccount@gmail.com   編輯

Create a password →   密碼   👁

繼續

已有帳戶？ 登入

Terms of use | Privacy policy

Check your email inbox for a verification email from OpenAI

Part 1-b. Generate An Essay with DaVinci

MediaTek DaVinci Platform

- DiVinvi (達哥) is a platform provided by MediaTek that hosts different AI models, including GPT-3.5 and GPT-4

DaVinci

1. Create a new chat

+ Chat

Search

No Data.

Plugins

Prompts

DVCx Library

Assistants

Model

GPT-3.5-Turbo-16k

2. Select a model

(Models has different pricing criteria)

Plugins

Select plugins used in chat

Conversation Style ⚪ Precise ⚪ Neutral 🔘 Creative

3. Select a Conversation Style

Precise - Always generates the same response for a given input

Neutral - Balances between precision and creativity

Creative - Encourages diversity and novelty

⚠️ Composing essays on DaVinci will consume your daily usage quota.

Remaining daily usage quota

4. Type your message here

⚡ Suggestion

Type a message or type "/" to select a prompt..

5. Send

More Actions   🟢 100% 🔔

Assistant   File   Prompt

+ Assistant

Search

DaVinci 達哥

Translation Pro

Doc Writing Expert

Visual Design Master - ...

GenAI Homework2 Assi...

GenAI Homework2 Assi...

25

Part 1-b. Generate An Essay with DaVinci: Demo

MediaTek DaVinci Tutorial

Click here to load DaVinci platform

DaVinci Chatbot Account Registration

- For **registered** students,

- You will receive the registration letter before 3/4 if you joined NTU Cool before noon on 3/1.

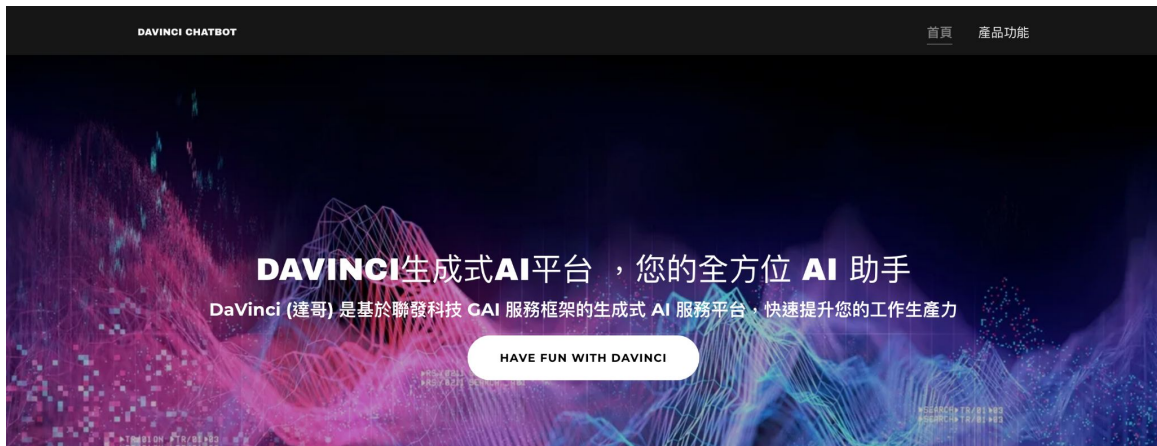- Otherwise, your account will be activated after 3/4.

⚠️ Please contact TAs if you haven't received a registration letter by 3/8.

Login

Your account should be set up through the email you provided.

If you do not reset the password **within 72 hours**, the account will be invalid.

This service is only open to registered students.

Part 2. Evaluate the Essays Using AI Models

Workflow

1. Generate an essay

Ask Language Models to write it
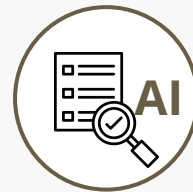
Write it on your own

Any other methods

Platforms
  a. OpenAI - ChatGPT
  b. MediaTek DaVinci（達哥）
  c. Others

2. Evaluate your Essay

We use generative AI as a TA to evaluate your essay automatically.

Generative AI TA [批改助教] Info：
1. Platform: MediaTek DaVinci
2. Model: GPT-4

3. Submit to NTU COOL

Essay Evaluation

In the past, the evaluation of essays are done by humans

The student
writes an essay

The teacher (human) evaluates the
essay **based on some criteria**

Obtain the evaluation result

Essay Evaluation

In this homework, the evaluation of essays are done by AI assistants

The student writes
an essay with AI

The **AI assistant** evaluates the
essay **based on some criteria**

Obtain the evaluation result

Essay Evaluation

- We employ evaluation assistants (批改助教) on MediaTek DaVinci platform

- We prepare **evaluation prompts** including evaluation criteria and steps

  - We combine an essay with the evaluation prompt and submit it to the assistant to obtain

    a score

Evaluation Prompt + Essay   **Evaluation Assistant**   **Assessment Result (Score)**

Essay Evaluation

Illustration of the evaluation by AI evaluation assistant

You are tasked with evaluating an article that presents the statement "Do you agree or disagree with the statement that Artificial Intelligence will eventually replace humans in most areas of work in the future world."
Your assignment involves assessing the article based on various criteria.
It is essential to thoroughly read and comprehend these instructions. Keep this document accessible during your review and refer to it as necessary.

Evaluation Criteria:
Ideas and Analysias (30%):
Evaluate the strength and depth of the article's ideas. Consider the analysis provided, the clarity of the main argument, and the overall coherence of the author's viewpoint.
(...)
Student's Essay:
<Essay to be evaluated>

Evaluation assistants

Final score: 7.5

Evaluation Settings in Evaluation Assistants

- Platform
  - MediaTek DaVinci Platform
- Model
  - GPT-4-turbo
- Hyperparameters
  - Temperature   0.5 (Neutral)

Evaluation Prompts

Due to the limited space, we put the full version of evaluation prompts below

- English

- Mandarin

Please contact TAs immediately if the links are not accessible.

Part 2. Evaluate the Essays on DaVinci Platform
Step 1: Install Evaluation Assistants

Evaluation Assistants

In Homework 2, you need two evaluation assistants:

- **GenAI Homework2 Assistant (English)**
  - for **English** essay evaluation
- **GenAI Homework2 Assistant (Mandarin)**
  - for **Mandarin** essay evaluation

⚠️ Please utilize the appropriate assistant to evaluate your essays. Otherwise, your evaluation results will be deemed **invalid and you will not receive any score**.

Chat

Chat

Q Search chat

▶ Folder_1

💬 Hello

Plugins   Prompts   DVCx Library   Assistants   ✕

3. Search the name of the assistant

🔍   GenAI Homework2 Assistant

All   2 assistants   Sort by

4. Click your target assistant   A - Z ▼

┌───┐
│ GenAI Homework2 Assistant (English)   Others ⭐ 0 (0) ⬇ 0 │
│ R12921120@NTU.EDU.TW   │
└───┘
An assistant for GenAI2024 Spring (a course led by Prof. Hung-Yi Lee in NTU) It aims to evaluate stu...

   GenAI Homework2 Assistant (Mandarin)   Others ⭐ 0 (0) ⬇ 0
   R12921120@NTU.EDU.TW

GenAI2024春季班的批改助教（由臺灣大學的李宏毅教授主講）其目的是評估學生的作業二（中文文章）並以分
數的形式提供反饋。

File   Prompt

File   + URL

Search DVCx

test.py

notes.txt

← Back   Assistant / **GenAI Homework2 Assistant (English)**

   GenAI Homework2 Assistant (English)   ┌─────────┐
                                                                  │ Install │
   R12921120@NTU.EDU.TW   └─────────┘

   Others ⭐ 0 (0) ⬇ 0 ↗ 🔗   **5. Install**

⋯ More Actions   41

Part 2. Evaluate the Essays on DaVinci Platform
Step 2: Use Assistants to Evaluate Essays

DaVinci

1. Create a new chat for each essay evaluation

+ Chat

Search

No Data.

Plugins

Prompts

DVCx Library

Assistants

2. Check the installed assistants

Assistant   File   Prompt

+ Assistant

Search

Model

GPT-3.5-Turbo-16k

Plugins

Select plugins used in chat

Conversation Style   ○ Precise   ○ Neutral   ● Creative

DaVinci 達哥

Translation Pro

Doc Writing Expert

Visual Design Master - ...

GenAI Homework2 Assi...

GenAI Homework2 Assi...

3. Click your target assistant

(GenAI Homework2 Assistant)

⚠ Ensure to open a **NEW** chat every time you re-evaluate the result

Suggestion

More Actions   100%

Type a message or type "/" to select a prompt...

4. Use the evaluation prompt to evaluate your essay

Evaluation prompt (English)

It's a prompt used for evaluating the English essays

Essay to be evaluated

Enter a value for Essay to be evaluated

5. Put your FULL essay here

Submit   **6. Click submit**

⚠️ Submitting the essay each time uses up your daily quota.

About 10 essay evaluations can be conducted in total each day.

Assistant : GenAI Homework2 Assistant (English) | Model : GPT-4-Turbo-8k | ⚙ 🗎 0 🗄

Example Essay1 (Bad):
"Artificial intelligence is bad. It will replace humans in the future. This is not good. We need jobs. Robots are not good. They will take our jobs. We need jobs to live. That is why I disagree."

Example Essay 2 (Average):
"Artificial Intelligence (AI) is becoming more prevalent in today's world. Its capabilities, like analyzing large data sets and automating repetitive tasks, make it a valuable asset in many industries. For instance, in healthcare, AI assists in diagnosing diseases with higher accuracy. However, its integration into the workforce raises concerns about job displacement.
Despite these advancements, AI lacks certain human qualities. It cannot replicate human empathy or creativity, which are crucial in professions like teaching or artistic endeavors. These roles require an understanding of human emotions and experiences that AI cannot provide. Therefore, while AI can support these fields, it cannot entirely replace the human element.
In conclusion, while AI's role is expanding, it is unlikely to completely replace humans in all job sectors. The balance between AI efficiency and human empathy is essential. Future advancements should focus on AI assisting humans, not replacing them, ensuring a harmonious coexistence between technology and humanity."

Student's Essay:
Title: The Role of Artificial Intelligence in the Future Workplace

In recent years, the advancement of Artificial Intelligence (AI) has sparked debates about its

🔄 Regenerate Response   ✦ Suggestion

⚡ Type a message or type "/" to select a prompt...   ➤   ⬇

45

argument by acknowledging the multifaceted impact of AI on the workforce. The evidence and reasoning provided are compelling and enhance the credibility of the article.

Organization: The essay is well-organized, with a logical flow of ideas and smooth transitions between paragraphs. The introduction sets up the argument, which is then explored through specific examples in the body paragraphs. The conclusion effectively summarizes the student's viewpoint, reinforcing the main argument. The essay adheres to the 300-word limit.

Language Use: The language used in the essay is clear, precise, and appropriate for the topic. Complex ideas, such as the irreplaceability of human creativity and empathy, are explained well. The writing is engaging and maintains the reader's interest throughout. Grammar and style are consistent with the standards of effective written communication.

Overall Assessment: The article effectively presents and supports the statement that AI will not replace humans in most areas of work. The argument is well-analyzed, developed, and organized, with appropriate language use. The student acknowledges the capabilities of AI while highlighting the indispensable human qualities that ensure the continued relevance of human workers in various professions.

Final score: 8.5/10

7. Assistant gives feedback and a final score for your essay

Regenerate Response   Suggestion

Type a message or type "/" to select a prompt...

46

argument by acknowledging the multifaceted impact of AI on the workforce. The evidence and reasoning provided are compelling and enhance the credibility of the article.

Organization: The essay is well-organized, with a logical flow of ideas and smooth transitions between paragraphs. The introduction sets up the argument, which is then explored through specific examples in the body paragraphs. The conclusion effectively summarizes the student's viewpoint, reinforcing the main argument. The essay adheres to the 300-word limit.

8. Export the result into a JSON file

Language Use: The language used in the essay is clear, precise, and appropriate for the topic. Complex ideas, such as the irreplaceability of human creativity and empathy, are explained well. The writing is engaging and maintains the reader's interest throughout. Grammar and style are consistent with the standards of effective written communication.

Overall Assessment: The article effectively presents and supports the statement that AI will not replace humans in most areas of work. The argument is well-analyzed, developed, and organized, with appropriate language use. The student acknowledges the capabilities of AI while highlighting the indispensable human qualities that ensure the continued relevance of human workers in various professions.

Final score: 8.5/10

⚠️ If there are good results, you can save them first to avoid any unexpected circumstances.

DaVinci

+ Chat

Search

Assessment of Artific...

More Actions   100%

Assistant   File   Prompt

+ Assistant

Search

DaVinci 達哥

Translation Pro

Doc Writing Expert

Visual Design Master - ...

GenAI Homework2 Assi...

GenAI Homework2 Assi...

Regenerate Response   Suggestion

Type a message or type "/" to select a prompt...

47

Part 2. Evaluate the Essays Using AI Models: Demo

Part 3. Submit to NTU Cool

Workflow

| 1. Generate an essay |

Ask Language Models to write it

Write it on your own

Any other methods

Platforms
a. OpenAI - ChatGPT
b. MediaTek DaVinci（達哥）
c. Others

| 2. Evaluate your Essay |

We use generative AI as a TA to evaluate your essay automatically.

AI

Generative AI TA [批改助教] Info :
1. Platform: MediaTek DaVinci
2. Model: GPT-4

| 3. Submit to NTU COOL |

Submission (NTU Cool)

- Complete the **Quiz HW2** on NTU Cool
    - Ensure to submit your evaluation result in **JSON** format
    - Submitting in the wrong format equals non-submission
- No restrictions on upload quota
- But only the most recent submission will be kept
- No late submission is allowed

⚠️ We will compare the files you submit with our backend data records to ensure they match. Therefore, please **do NOT** manually edit your JSON file.

Quick Demonstration

Quick Demonstration

1. Use any platform to generate an essay. (Here we use ChatGPT as example)
2. Evaluate the essay by GenAI Homework2 Assistant. (on DaVinci Platform)
3. If you are satisfied with the result, download the json file and submit it.

Do you agree or disagree with the statement that Artificial Intelligence will eventually replace humans in most areas of work in the future world.

Try different prompt here!

ChatGPT

The impact of Artificial Intelligence (AI) on the future workforce is complex. While AI may automate some tasks, leading to job displacement in certain industries, it's unlikely to completely replace humans in most areas of work. Many jobs require qualities like creativity and social interaction, which are challenging for AI to replicate. Instead of focusing solely on replacement, some suggest a future where AI augments human capabilities, leading to collaboration rather than outright displacement. Overall, the future of work will likely involve a mix of automation and new opportunities, with varying impacts across different sectors.

Evaluation prompt (English)

It's a prompt used for evaluating the English essays

Essay to be evaluated

capabilities, leading to collaboration rather than outright displacement. Overall, the future of work will likely involve a mix of automation and new opportunities, with varying impacts across different sectors.

Submit

Submit json file

Grading Policy

Point Distribution

- Homework 2 accounts for 10 points in total

- (3pts) NTU Cool Submission
    - You need to complete the entire quiz to get full credit

- (7pts) Assessment of two essays
    - 3.5 points for each essay

Grading — English Essay

- **Essay Score**   **Homework Score**
- 9.5 — 10.0   +3.5 pts
- 8.5 — 9.4   +3.0 pts
- 8.0 — 8.4   +2.5 pts
- 7.5 — 7.9   +2.0 pts
- 7.0 — 7.4   +1.5 pts
- 6.0 — 6.9   +1.0 pts
- Otherwise   +0.0 pts

Grading – Mandarin Essay

- **Essay Score**   **Homework Score**
- Rank A+   +3.5 pts
- Rank A   +3.0 pts
- Rank B+   +2.5 pts
- Rank B   +2.0 pts
- Rank C+   +1.5 pts
- Rank C   +1.0 pts
- Otherwise   +0.0 pts

Important Dates

Important Dates

- Deadline for Submission (NTU Cool)

    2024/03/14 23:59:59 (UTC+8)

- Grading Release Date

    2024/03/28 23:59:59 (UTC+8)

Regulations

Regulations

- Plagiarism in any form is prohibited.
- Do NOT share your prompts and evaluation results (JSON files) with others.
- Do NOT submit the JSON files that are not obtained using your Davinci account
- Do NOT attempt to manually edit your JSON files.
- 第一次違反以上規定，**該作業0分，學期總成績再乘以0.9**
- 第二次違反以上規定，**學期成績F**
- Prof. Lee & the TAs preserve the rights to change the rules & grades.

If You Have Any Questions

- NTU Cool [HW2 Forum](#)
    - Recommended if no assignment answers or privacy issues are involved in your questions
    - Prioritize answering questions on the homework forum
- Email: ntu-gen-ai-2024-spring-ta@googlegroups.com
    - Title should start with [GenAI 2024 Spring HW2]
    - Email with the wrong title will be moved to trash automatically
- TA Hours
    - Time: **3/1、3/8 (16:30 - 17:20)**
    - Location: **綜合大講堂**

Reminders for using DaVinci platform

There is a daily usage quota limitation on the DaVinci platform

- Each time you use the platform to generate or evaluate something, it consumes your quota

  - For evaluating essays using the evaluation assistant: Approximately **10** essay evaluations can be conducted in total per day (if you do not use DaVinci to do anything else)

  - For generating essays: Approximately **500** essays can be generated using GPT-3.5 model. Using the GPT-4 model will incur higher expenses. (20x expensive than GPT-3.5)