

TWLLM Tutor: Revolutionizing Taiwanese Secondary Education with Large Language Model



Jia-Wei Liao



Ji-Jia Wu



Kun-Hsiang Lin



Kang-Yang Huang

National Taiwan University, Taipei, Taiwan
<https://github.com/jwliao1209/TWLLM-Tutor>



Outline

- Introduction
- Methods
- Results
- Conclusion

Outline

- Introduction
- Methods
- Results
- Conclusion

Education



Rural Education



Education for Disadvantaged

Taiwan Culture

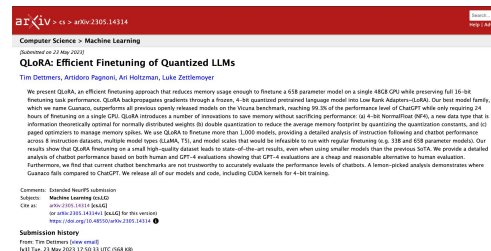


News



Chinese Language Test

Artificial Intelligence



State-of-the-art Technology



LLM

Introduction to Datasets

- General scholastic ability test social studies (GSAT, 學測): 1434 (#questions)
- Social studies question bank with answer explanation
 - Geography: 322 / History: 9058 / Civics: 2053 (#questions)

財團法人大學入學考試中心基金會

112學年度學科能力測驗試題

社會考科

請於考試開始鈴響起，在答題卷簽名欄位以正楷簽全名

— 作答注意事項 —

考試時間：110分鐘

作答方式：

- 選擇題用 2B 鉛筆在「答題卷」上作答；更正時以橡皮擦擦狀，切勿使用修正帶（液）。
- 除題目另有規定外，非選擇題用筆尖較細之黑色墨水的筆在「答題卷」上作答；更正時，可以使用修正帶（液）。
- 考生須依上述規定劃記或作答，若未依規定而導致答案難以辨識或評閱時，恐將影響成績。
- 答題卷每人一張，不得要求增補。

選擇題計分方式：

- 單選題：每題有 n 個選項，其中只有一個是正確或最適當的選項。答對者，得該題的分數；答錯、未作答或劃記多於一個選項者，該題以零分計算。

1 - 1 文學與哲學

單選題

題號：0101-00008

085306-0101-00008

難易度：易

出處：精選試題

印度古典時期，在種族與語言文化的發展上呈現多元複雜的特點。請問：在此時期曾經侵入印度地區的族群在先後順序上，下列選項何者正確？

(A)阿利安人→希臘人→大月氏人 (B)阿利安人→大月氏人→希臘人 (C)阿利安人→大月氏人→波斯人 (D)阿利安人→波斯人→大月氏人

答案：(A)

題號：0101-00009

085306-0101-00009

難易度：易

出處：精選試題

西元前四世紀，☐進入印度河流域，助長當時印度內部的動亂。他們在印度停留的時間雖然有限，但卻成為東西方文明之間的首次大規模接觸。請問：☐應為下列哪一族群？

(A)拉齊普特人 (B)大月氏人 (C)希臘人 (D)匈奴人

答案：(C)

解析：希臘亞歷山大的軍隊於西元前 326 年進入印度河流域。

Introduction to Our Approaches

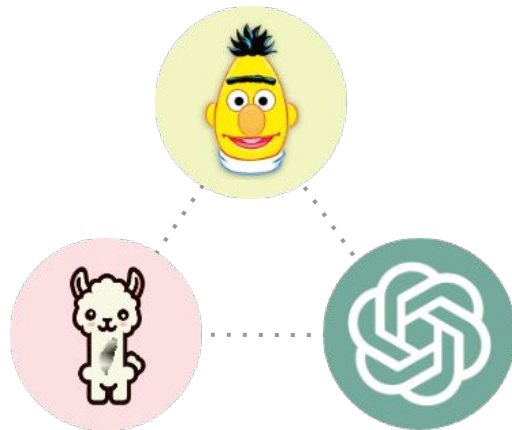
- **BERT-based Model**

- Multiple choice
- Vision

- **Taiwan LLM**

- Prompt engineering
- QLoRA / LoftQ
- Instruction tuning
- Multiple choice (classification)
- Vision

- **ChatGPT**



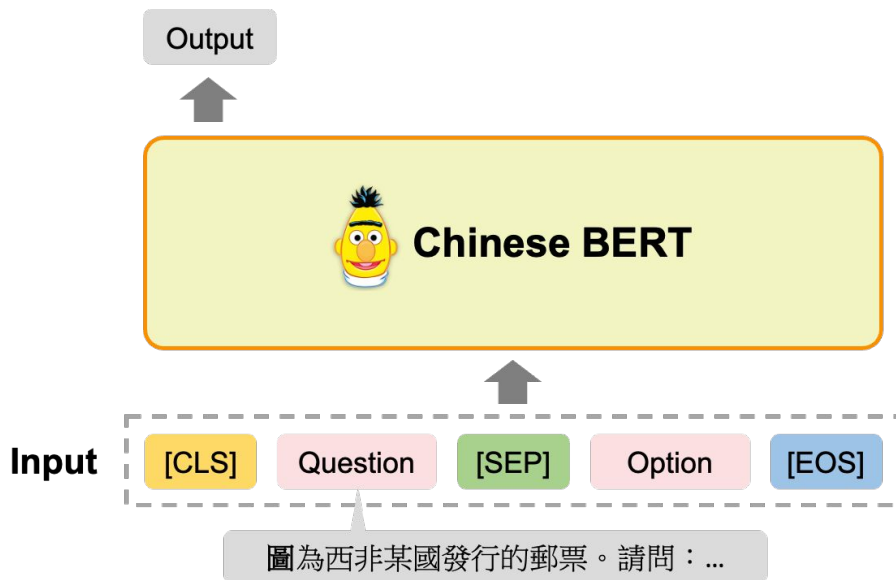
Outline

- Introduction
- **Methods**
- Results
- Conclusion



Architecture (I): BERT for Multiple Choice

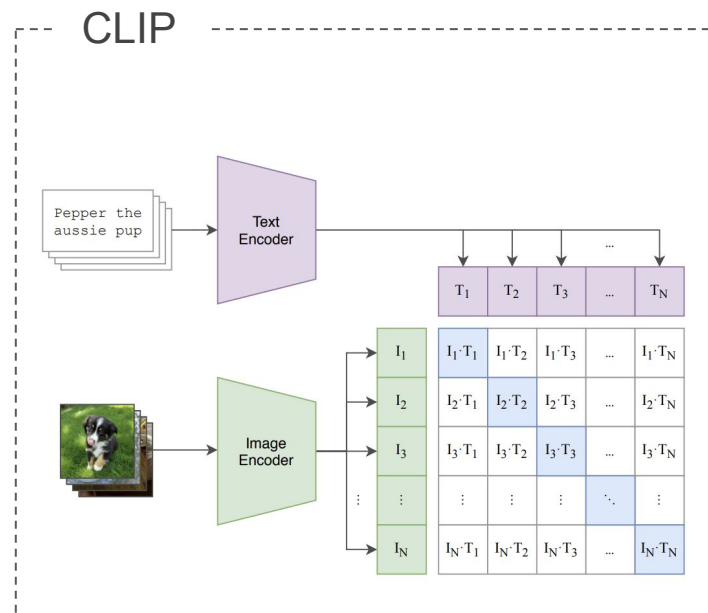
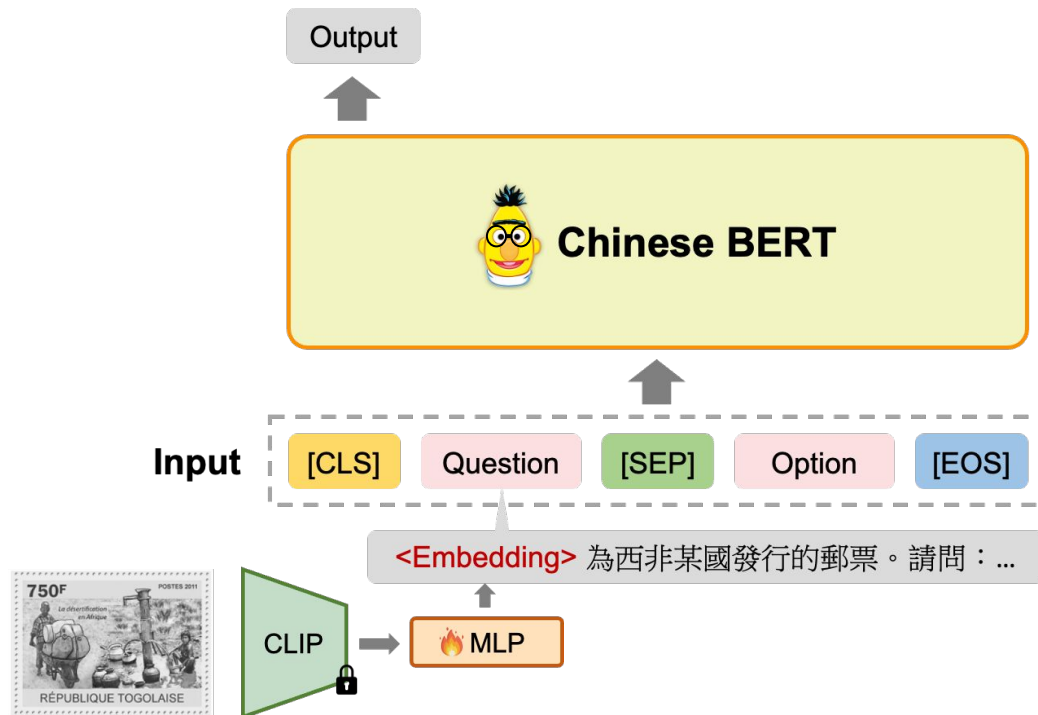
- Pre-trained model
 - chinese-base-bert [\[CKIP Lab, 2020\]](#)
 - chinese-roberta-wwm-ext [\[Cui, et al., 2021\]](#)



Architecture (II): Vision BERT for Multiple Choice



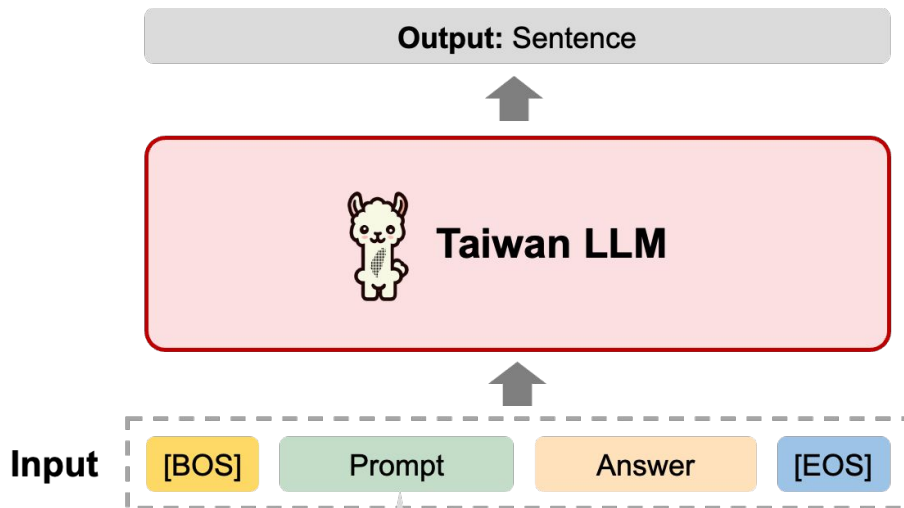
- Utilize pre-trained CLIP [Radford, et al., 2021] for zero-shot transformation



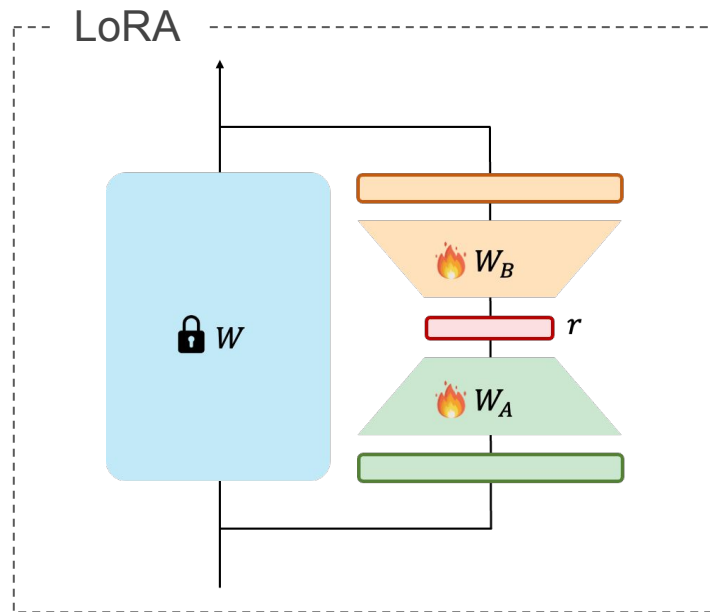
Architecture (III): Taiwan LLM Instruction Tuning



- Pre-trained model: Taiwan-LLM-7B-v2.0.1-chat [Lin, et al., 2023]



<Prompt Prefix> [USER]: 就我國刑事訴訟制度而言，高等法院再經審理後做成的判決，其判決理由是在實踐什麼原則？
A. 審檢分立 \n B. 罪刑法定 \n C. 從舊從輕 \n D. 無罪推定 \n
[ASSISTANT]: 正確答案為



Architecture (III): Taiwan LLM Instruction Tuning



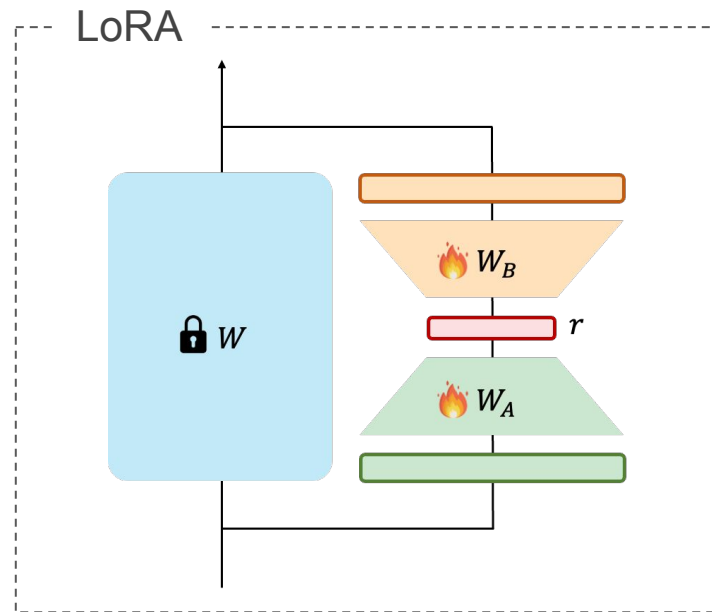
$$\min_{Q,A,B} \|W - Q - AB^T\|_F$$

Algorithm 1 LoftQ

input Pre-trained weight W , target rank r , N -bit quantization function $q_N(\cdot)$, alternating step T

- 1: Initialize $A_0 \leftarrow 0, B_0 \leftarrow 0$
- 2: **for** $t = 1$ to T **do**
- 3: Obtain quantized weight $Q_t \leftarrow q_N(W - A_{t-1}B_{t-1}^T)$
- 4: Obtain low-rank approximation $A_t, B_t \leftarrow \text{SVD}(W - Q_t)$ by (9)
- 5: **end for**

output Q_T, A_T, B_T





Architecture (IV): Taiwan LLM Multiple Choice

Output: 0 (A), 1 (B), 2 (C), 3 (D)

$$\min_{Q,A,B} \|W - Q - AB^T\|_F$$

Layer

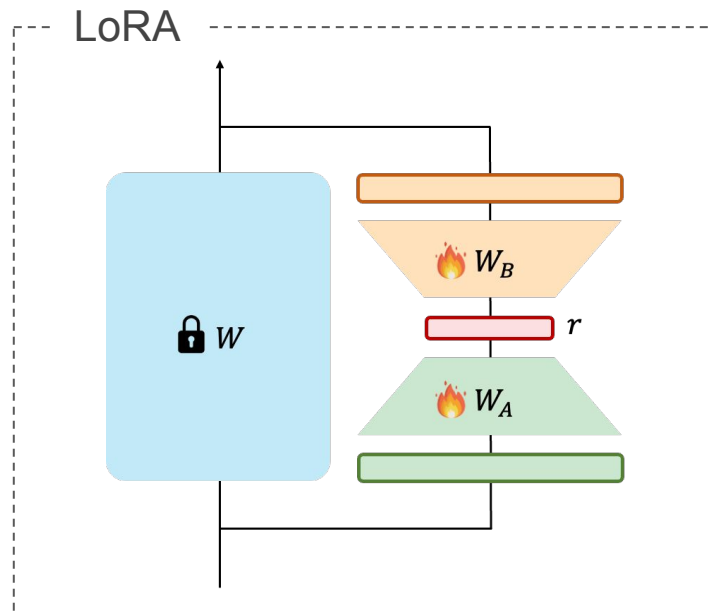
Algorithm 1 LoftQ

input Pre-trained weight W , target rank r , N -bit quantization function $q_N(\cdot)$, alternating step T

- 1: Initialize $A_0 \leftarrow 0, B_0 \leftarrow 0$
- 2: **for** $t = 1$ to T **do**
- 3: Obtain quantized weight $Q_t \leftarrow q_N(W - A_{t-1}B_{t-1}^T)$
- 4: Obtain low-rank approximation $A_t, B_t \leftarrow \text{SVD}(W - Q_t)$ by (9)
- 5: **end for**

output Q_T, A_T, B_T

<Prompt Prefix> [USER]: 就我國刑事訴訟制度而言，高等法院再經審理後做成的判決，其判決理由是在實踐什麼原則？
A. 審檢分立 \n B. 罪刑法定 \n C. 從舊從輕 \n D. 無罪推定 \n
[ASSISTANT]: 正確答案為



New Optimizer: Lion

- Symbolic discovery of optimization algorithms [\[Chen, et al., 2023\]](#)

AdamW

$$\mathbf{m}_t = \beta_1 \mathbf{m}_{t-1} + (1 - \beta_1) \mathbf{g}_t$$

$$\mathbf{v}_t = \beta_2 \mathbf{v}_{t-1} + (1 - \beta_2) \mathbf{g}_t^2$$

$$\hat{\mathbf{m}}_t = \frac{\mathbf{m}_t}{1 - \beta_1^t}$$

$$\hat{\mathbf{v}}_t = \frac{\mathbf{v}_t}{1 - \beta_2^t}$$

$$\mathbf{u}_t = \frac{\hat{\mathbf{m}}_t}{\sqrt{\hat{\mathbf{v}}_t} + \epsilon}$$

$$\boldsymbol{\theta}_t = \boldsymbol{\theta}_{t-1} - \eta_t (\mathbf{u}_t + \lambda_t \boldsymbol{\theta}_{t-1})$$



Lion

$$\mathbf{u}_t = \text{sign}(\beta_1 \mathbf{m}_{t-1} + (1 - \beta_1) \mathbf{g}_t)$$

$$\boldsymbol{\theta}_t = \boldsymbol{\theta}_{t-1} - \eta_t (\mathbf{u}_t + \lambda_t \boldsymbol{\theta}_{t-1})$$

$$\mathbf{m}_t = \beta_2 \mathbf{m}_{t-1} + (1 - \beta_2) \mathbf{g}_t$$



Outline


- Introduction
- Methods
- **Results**
- Conclusion

BERT for Multiple Choice



Multiple Choice with Image Results

- Testing dataset: 108-112 GSAT

Training data	Model	Accuracy
83-107 GSAT	BERT	0.3351
83-107 GSAT	Vision-BERT	0.3514
83-107 GSAT + Question bank	BERT	0.3892
 83-107 GSAT + Question bank	Vision-BERT	0.4162

Taiwan LLM Tuning



Prompt Designing

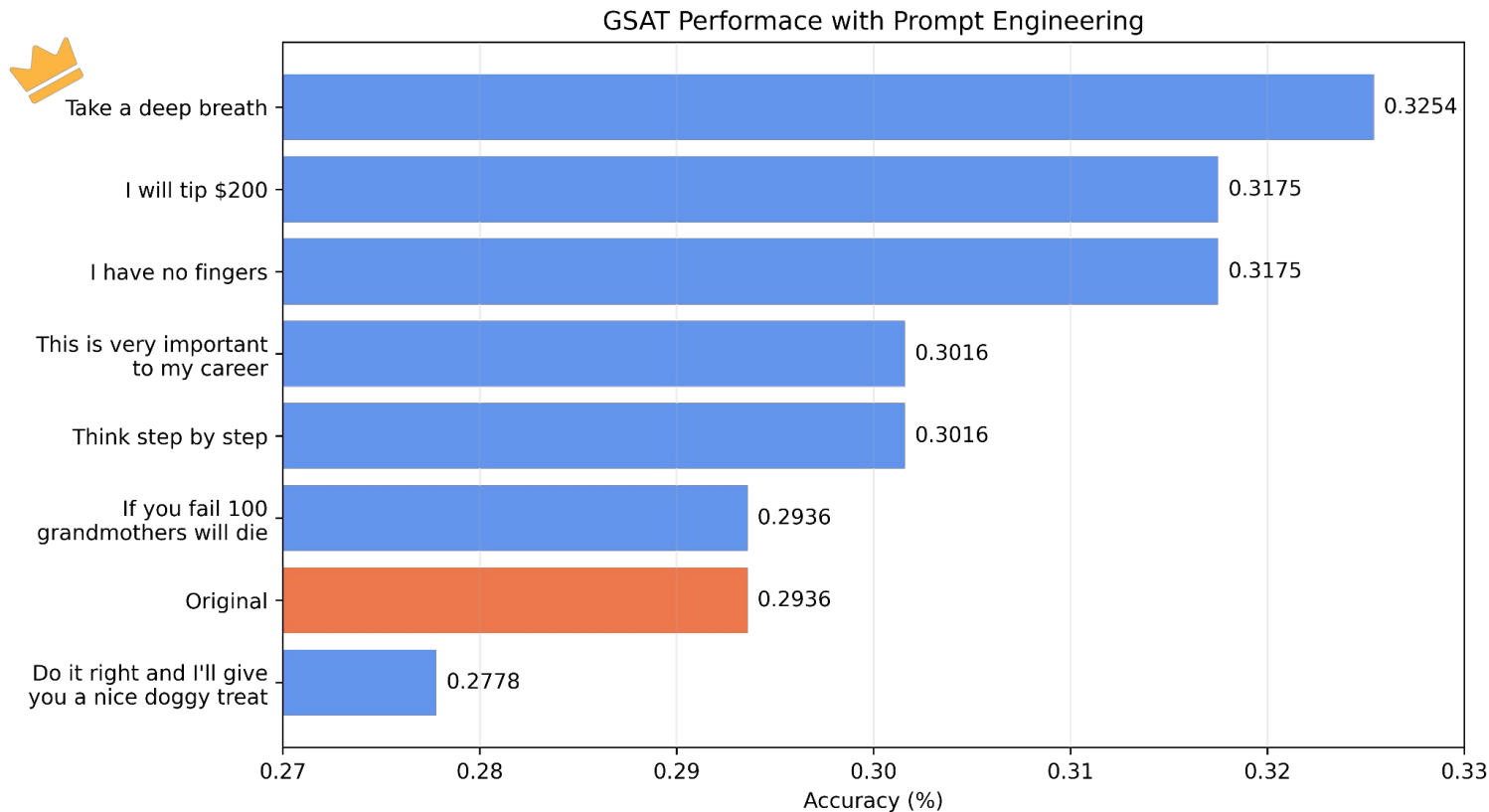
Question

<Prompt Prefix> 你是一名人工智慧家教，以下的題目為高中社會科，請根據題目提供使用者正確答案的選項 A 或 B 或 C 或 D。USER: **<Question and Options>**
ASSISTANT: 正確答案為

Answer

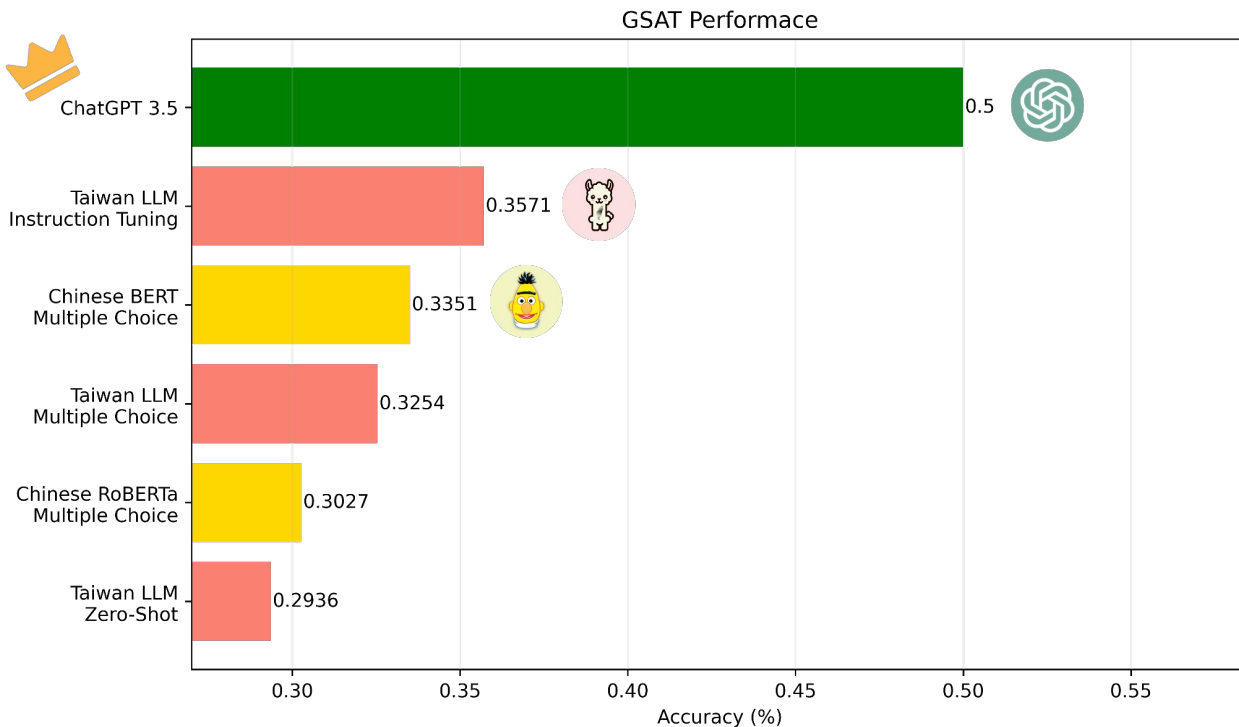
<Correct Option> 原因: **<Answer Explanation>**

Prompt Engineering Results



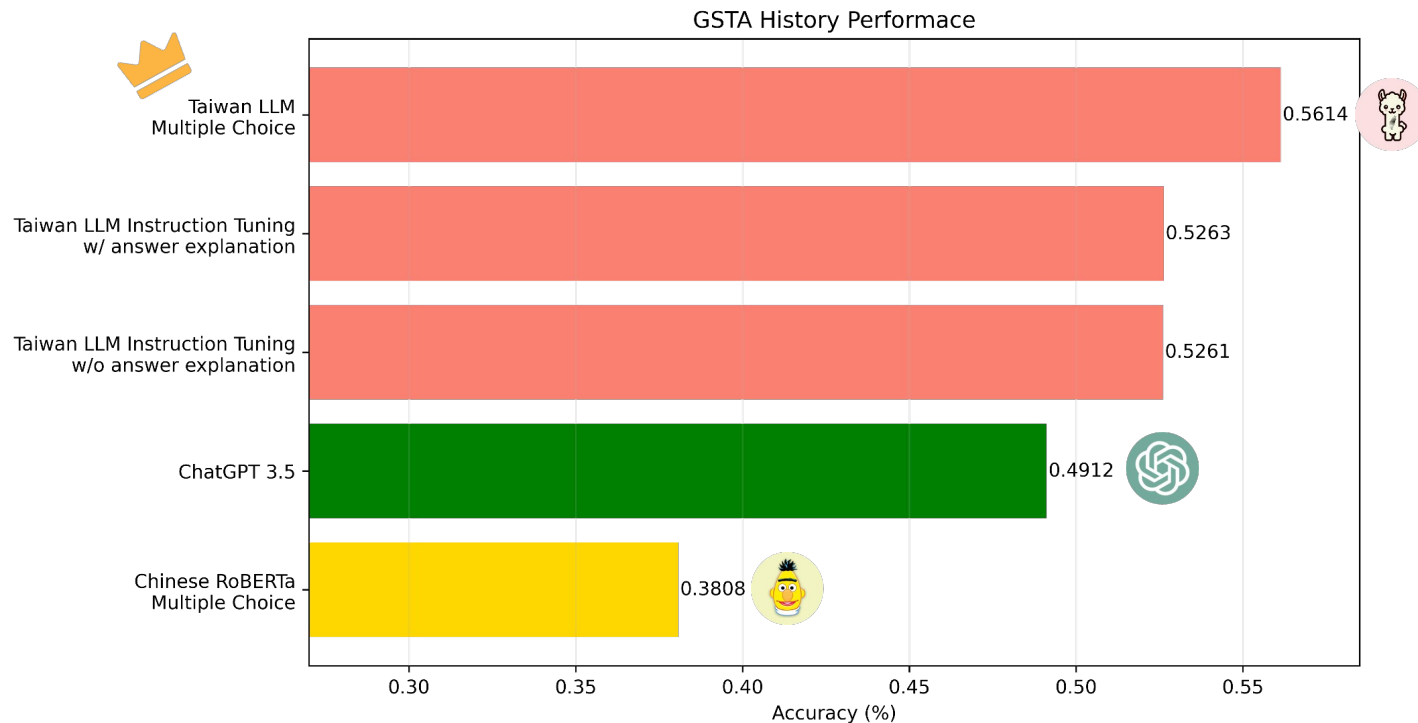
GSAT Results

- 83-109 GSAT for training and 110-112 GSAT for testing



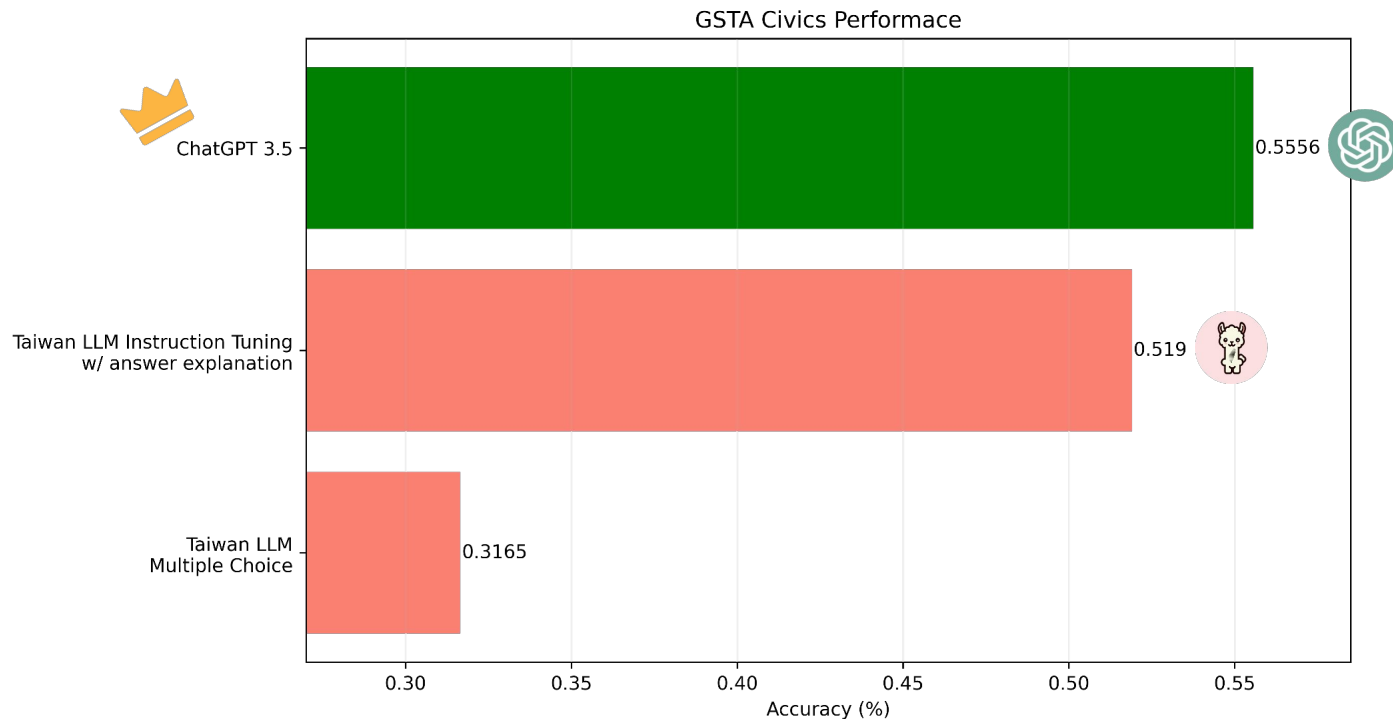
History Results

- History question bank for training and 110-112 GSAT for testing



Civics Results

- Civics question bank for training and 110-112 GSAT for testing



Case Study



你是一名人工智慧家教，以下的題目為高中社會科，請根據題目提供使用者正確答案的選項 A 或 B 或 C 或 D。

十九世紀前期，張三從寧波來艋舺採購樟腦，計畫運回浙江、上海銷售，同時也運來綢緞、布帛等貨品來臺銷售。請問：張三最可能委託下列哪個組織運送這批貨物？ A.東印度公司 B.北郊 C.怡和洋行 D.三井會社



B.北郊 原因：當時臺北港的北郊商人常兼職貿易商，對於貨物的運送和買賣多半有涉獵。



C. 怡和洋行 原因：在十九世紀前期，張三要將樟腦、綢緞、布帛等貨品運送回浙江和上海銷售，最可能委託的組織是怡和洋行。怡和洋行是一家在19世紀初期在華南地區經營貿易的外國商行，主要從中國向其他國家運送貨物。因此，怡和洋行是最可能協助張三運送這批貨物的組織。其他選項，如A. 東印度公司、B. 北郊、D. 三井會社，都不太可能參與這個情境中的貨物運送。



Case Study



你是一名人工智慧家教，以下的題目為高中社會科，請根據題目提供使用者正確答案的選項 A 或 B 或 C 或 D。

學者指出：北魏實行均田法的背景，第一，因戰亂的傷亡及遷徙，留下大片無主荒田；第二，許多農民為躲避戰禍，逃離農村，流落城市，成為不事生產的游民；第三，農民投靠大戶士族，「五十家方為一戶」，只向其宗主貢獻，不必向政府納稅。均田法的實施是結合戶籍的清查及管理，透過計口授田，使農業勞動力與耕地達到合理配合，以恢復生產，建立小農社會。根據上述，學者認為北魏均田法實施的主要目的最可能是：A.限制土地兼併 B.保全士族特權 C.重建編戶稅基 D.推行漢化運動



C.重建編戶稅基 原因：北魏均田法的實施，使農民戶籍清查並管理，透過計口授田，使農業勞動力與耕地達到合理配合，以恢復生產，重建編戶稅基。
\\n(A)與資料無關。
\\n(B)與資料無關。
\\n(D)與資料無關。



A. 限制土地兼併 原因：根據提供的資訊，北魏實行均田法的背景包括大片無主荒田、許多農民流落城市成為游民、農民投靠大戶士族等情況。均田法的實施主要目的是限制土地兼併，通過計口授田，確保農業勞動力和土地之間的合理配合，防止土地被少數大戶士族壟斷，促進農業生產，建立小農社會。因此，A. 限制土地兼併是最合適的答案。



Case Study



你是一名人工智慧家教，以下的題目為高中社會科，請根據題目提供使用者正確答案的選項 A 或 B 或 C 或 D。

1949年10月，共軍集中船隻對金門發動攻擊，遭守軍擊退。1950年3月，共軍又集結大批兵力，用類似「萬船齊放戰術」進攻海南島，迫使當地國軍撤守；5月底更攻占廣東萬山群島。共軍原本預定隨後再度對金門發動攻擊，但沒有多久就取消這項計畫。最可能的原因是：A. 共軍軍備及船隻有限，已經無法再負擔長期作戰 B. 共軍顧忌金門兵力強大，寧可保持戰果故而避戰 C. 韓戰爆發，美國宣布派遣第七艦隊協助臺灣防務 D. 臺灣實施戒嚴，共軍乃轉以心戰策略來瓦解國軍



C. 韓戰爆發，美國宣布派遣第七艦隊協助臺灣防務 原因：韓戰爆發後，美國宣布派遣第七艦隊協助臺灣防務，因此共軍暫緩對金門的攻擊。



A. 共軍軍備及船隻有限，已經無法再負擔長期作戰 原因：1949年10月，共軍對金門發動攻擊，但遭到守軍擊退。隨後，共軍又進攻海南島並攻占廣東萬山群島。然而，最可能的原因取消對金門的再次攻擊計畫是因為共軍的軍備和船隻有限，已經難以再負擔長期作戰，尤其在面對金門守軍強大的情況下。共軍可能決定將資源用於其他地區或其他策略，因為攻擊金門可能會導致更大的損失。其他選項如B. 共軍顧忌金門兵力強大、C. 韓戰爆發，...



Outline

- Motivation
- Methods
- Results
- Conclusion

Conclusion

- Create an academic dataset for AI applications in education.
- We surveyed several LLM tuning methods to achieve competitive performance comparing with ChatGPT.
- We took the first step to integrate visual data with our model's input, enabling it to answer questions not only based on language data but also on visual information.

Future Work

- Search for more potential methods for lightweight fine-tuning (e.g., QLoRa) to address computational resource limitations.
- Investigate more appropriate approaches to integrate image-related data, such as figures and tables, into the model input in order to improve overall accuracy since questions often consist figures or tables.
- Explore methods like reinforcement learning from human feedback (RLHF) to enhance the quality of the provided detailed explanations.
- Collaborate with Junyi (均一) Academy.



Reference

1. CKIP Lab, <https://github.com/ckiplab/ckip-transformers>, 2020.
2. Cui, Yiming, et al. "Pre-training with whole word masking for chinese bert." IEEE/ACM Transactions on Audio, Speech, and Language Processing 29 (2021): 3504-3514.
3. Radford, Alec, et al. "Learning transferable visual models from natural language supervision." International conference on machine learning. PMLR, 2021.
4. Lin, Yen-Ting, and Yun-Nung Chen. "Taiwan LLM: Bridging the Linguistic Divide with a Culturally Aligned Language Model." arXiv preprint arXiv:2311.17487 (2023).
5. Chen, Xiangning, et al. "Symbolic discovery of optimization algorithms." arXiv preprint arXiv:2302.06675 (2023).

Thank you

願有一天台灣所有的孩子都能快樂學習
勇敢追尋屬於自己的夢想