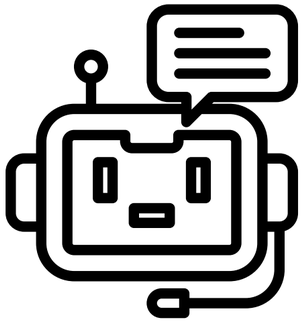
LLM Classification Finetuning

第87組 李宥萱 吳定霖 吳宗樺

Presentation Video Link:
https://drive.google.com/file/d/1A8whU4EYb9M39sdzyfvjDEdaSVYHQbLy/view?usp=sharing

目錄

任務介紹

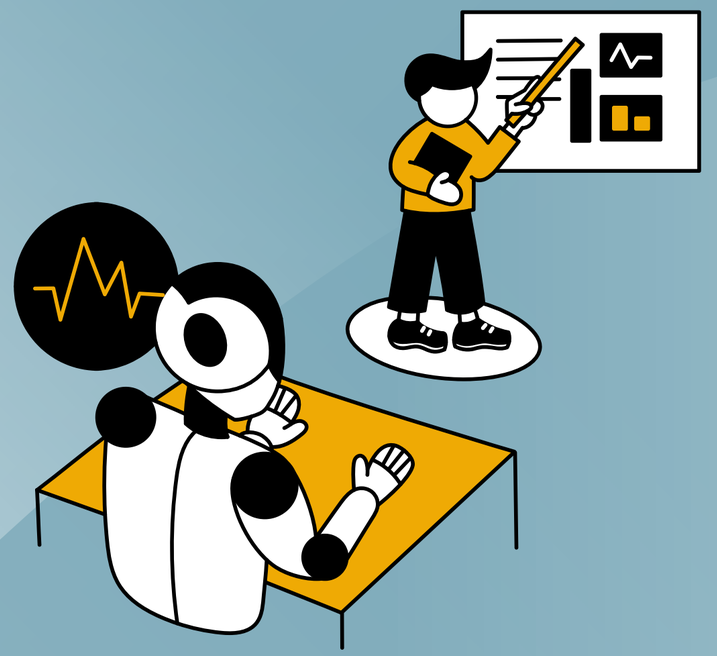


此任務目的是預測使用者更偏好哪一個大型語言模型所回答的答案，幫助縮短 LLM 能力與人類偏好之間的差距。

與Reinforcement Learning from Human Feedback的概念密切相關，將有助於改善聊天機器人與人類的互動。

Reinforcement Learning from Human Feedback

- Reinforcement Learning
  - 透過設定要演算法達成的目標，然後「獎勵模型」會根據演算法嘗試的結果給予回饋值，以取得最大化的預期利益
- Human Feedback
  - 根據人類回饋訓練「獎勵模型」
  - 不僅是根據固定的規則學習，也通過人類的指導來理解更複雜的部分，使得機器能夠更貼近人類的期望和行為方式。



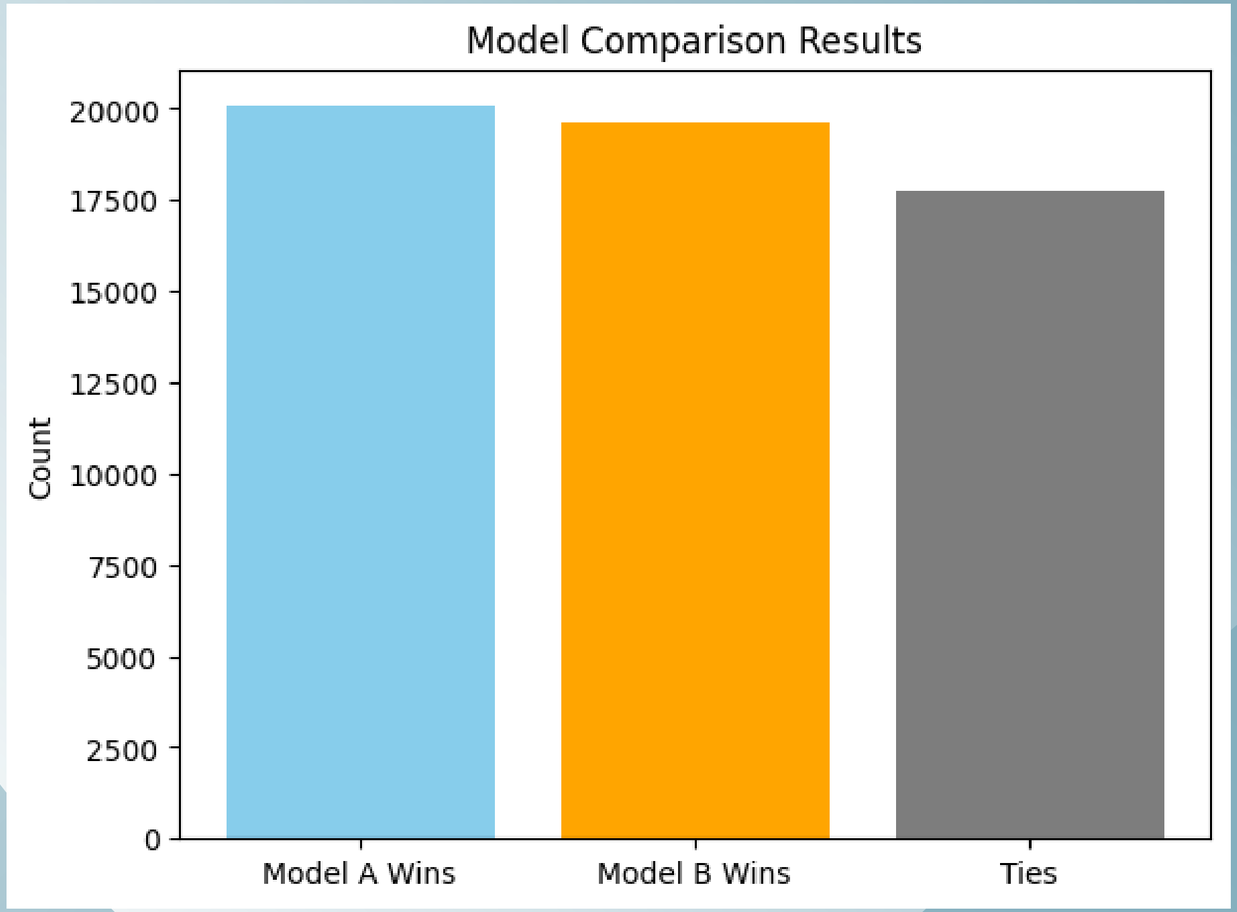
資料集介紹

- 資料集由 ChatBot Arena 的使用者互動資料組成。在每次使用者互動中，會向兩個不同的大型語言模型提供一個或多個prompt，然後指出哪個模型給出的回應更令人滿意。
- 訓練資料包含 55,000 筆資料，而測試集約為 25,000 筆。

欄位	描述
id	A unique identifier for the row.
model_[a/b]	The identity of model_[a/b]. Included in train.csv but not test.csv.
prompt	The prompt that was given as an input (to both models).
response_[a/b]	The response from model_[a/b] to the given prompt.
winner_model_[a/b/tie]	Binary columns marking the judge's selection. The ground truth target column.

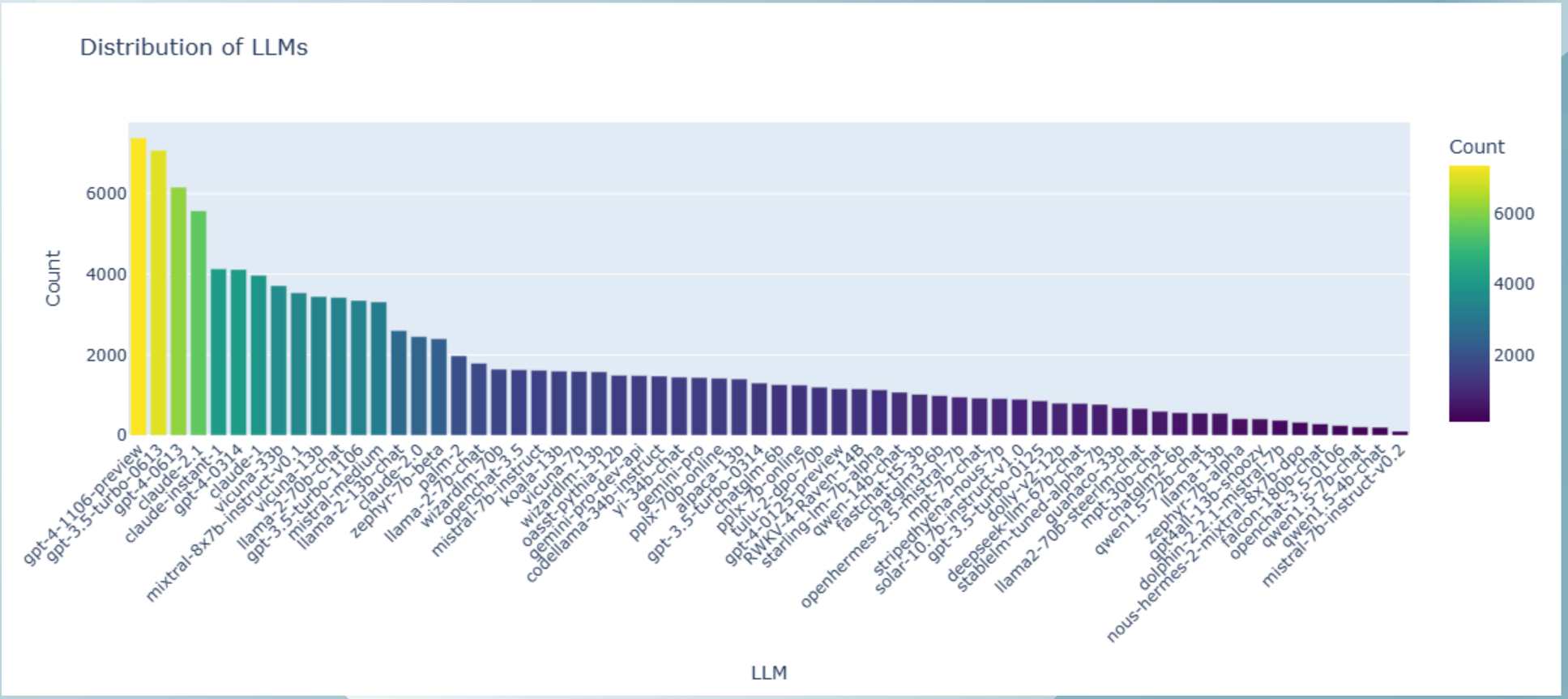
- 輸入: prompt，response_[a/b]
- 輸出: a 模型回應較好的機率、b 模型回應較好的機率、兩模型回應一樣好的機率 (三個機率值總和為 1)
- 評估方式: 三個預測機率值與正確答案的 Log Loss

EDA



- 平手的情形共出現17761次

EDA

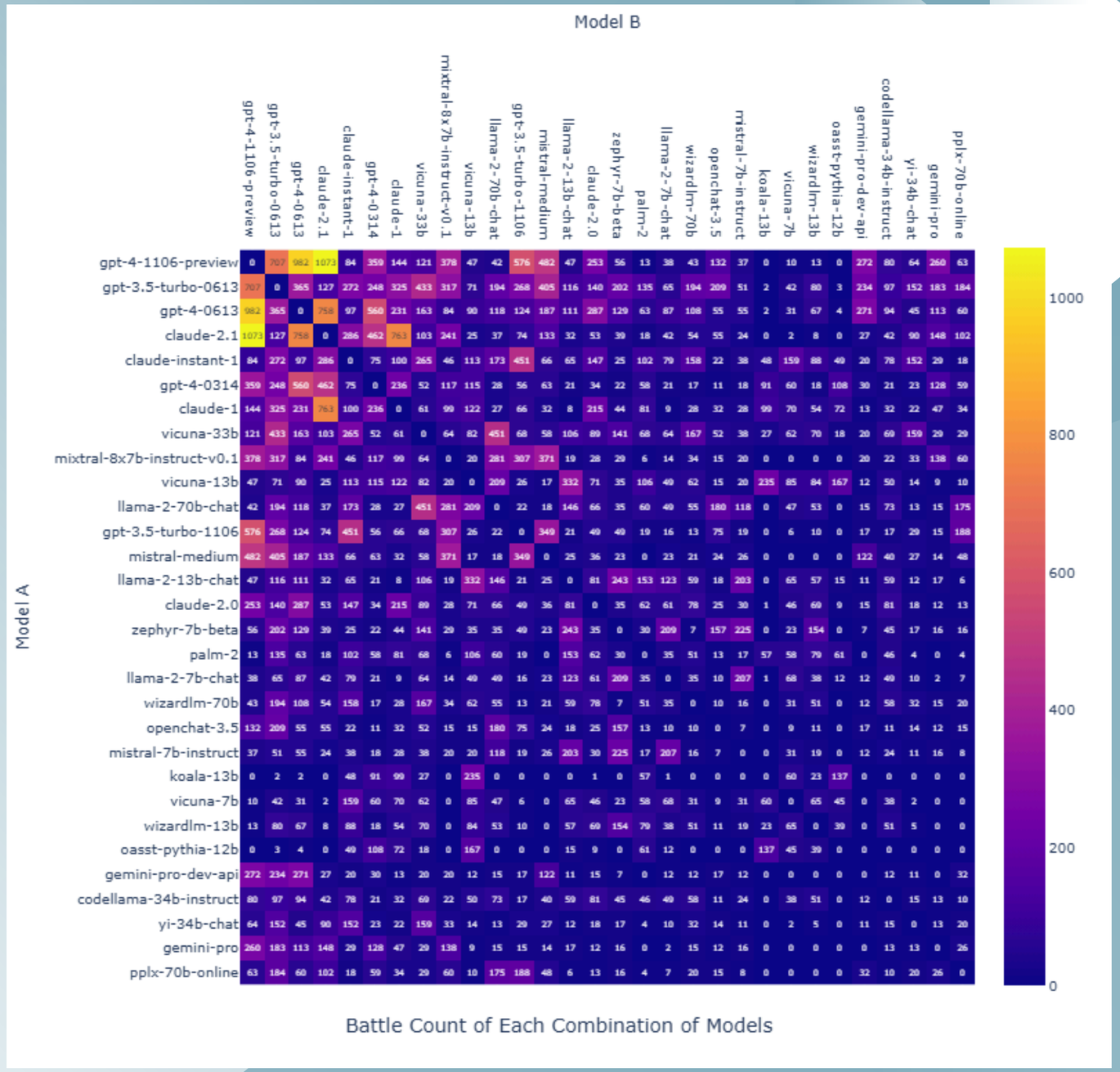


- 共64種模型
- 最常被做比較的模型前三名分別為gpt-4-1106-preview, claude-2.1, claude-1

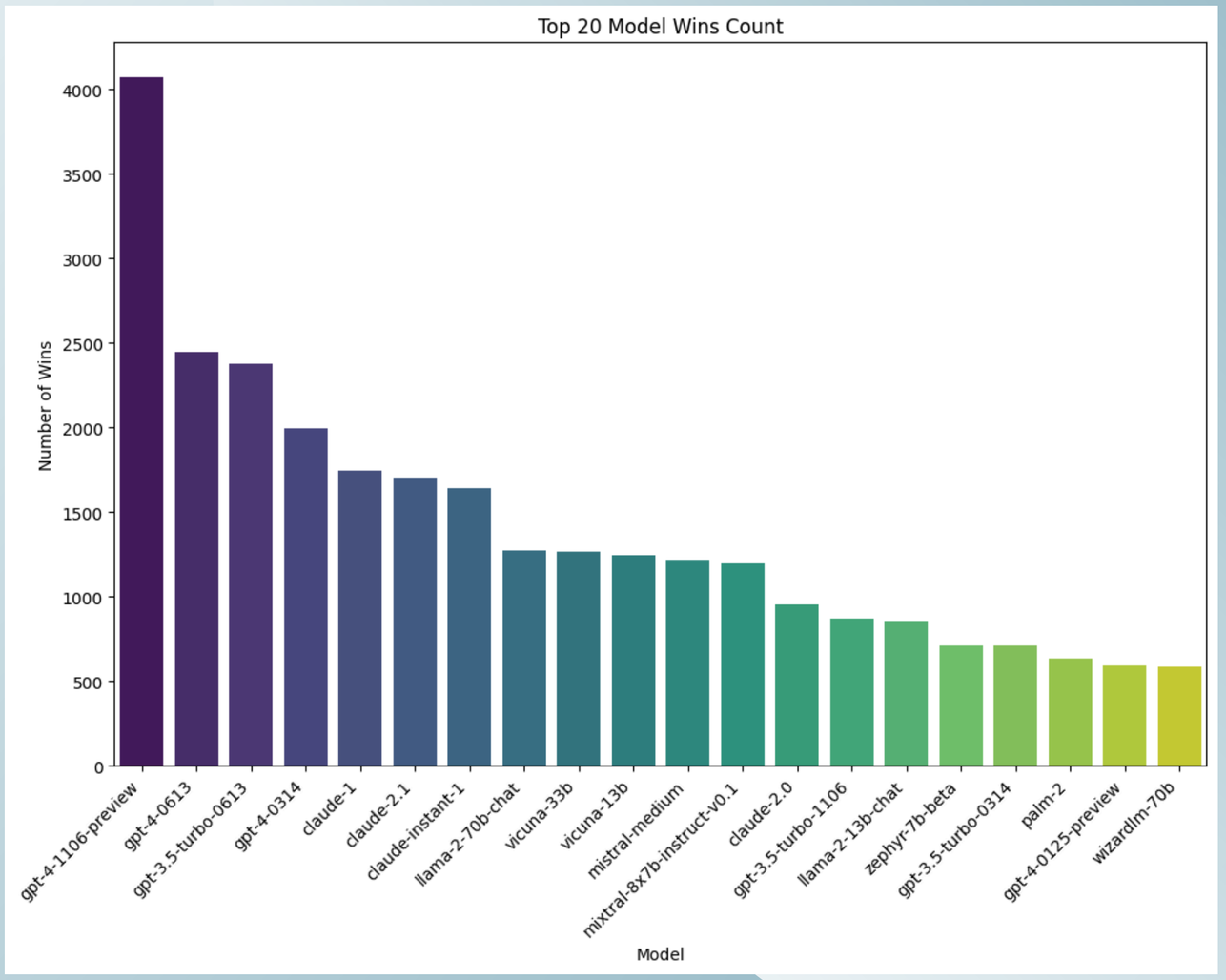
EDA

最常被兩兩比較的模型：

1. gpt-4-1106-preview V.S. claude-2.1
2. gpt-4-1106-preview V.S. gpt-4-0613
3. claude-1 V.S. claude-2.1



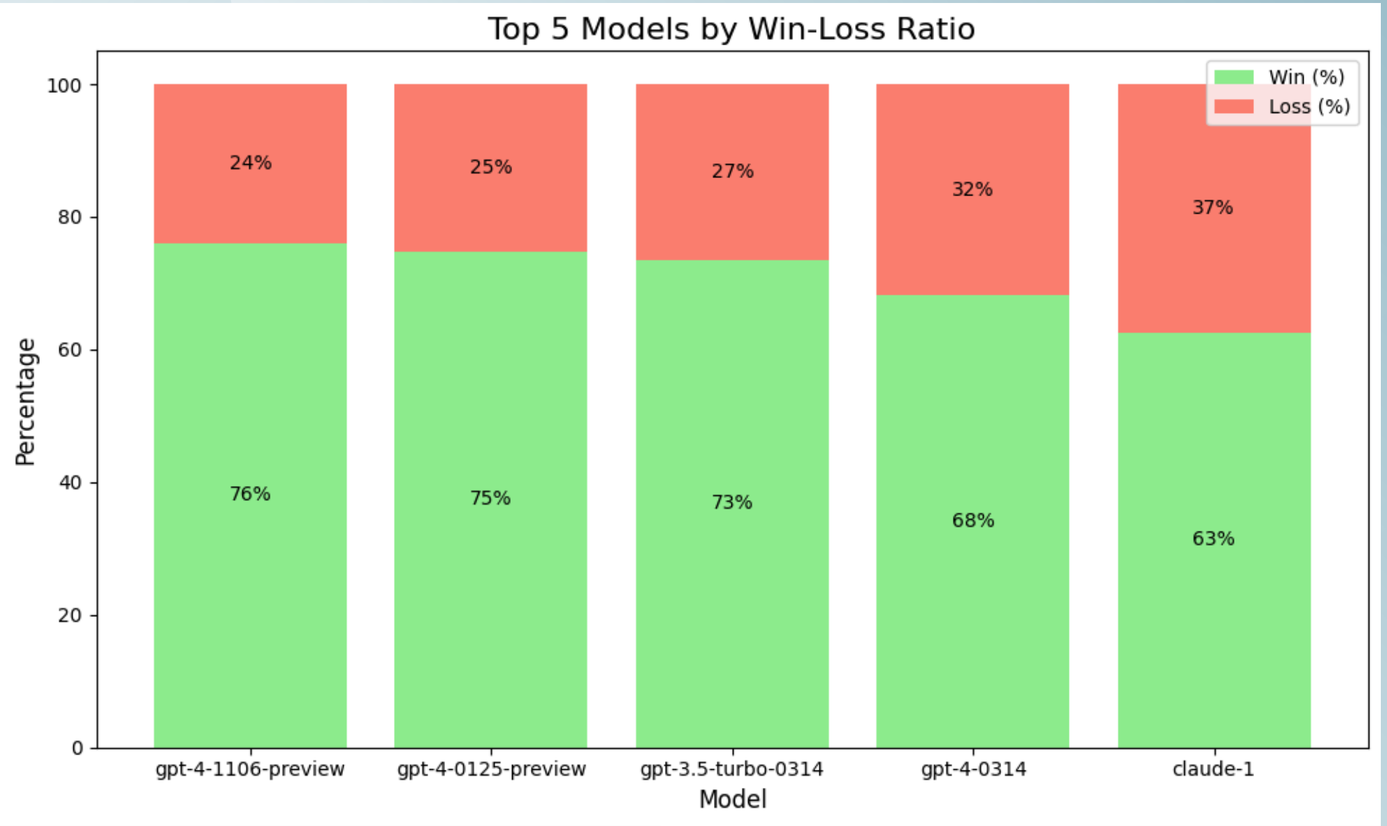
EDA



模型贏的次數最多的前三名：

1. gpt-4-1106-preview
2. gpt-4-0613
3. gpt-3.5-turbo-0613

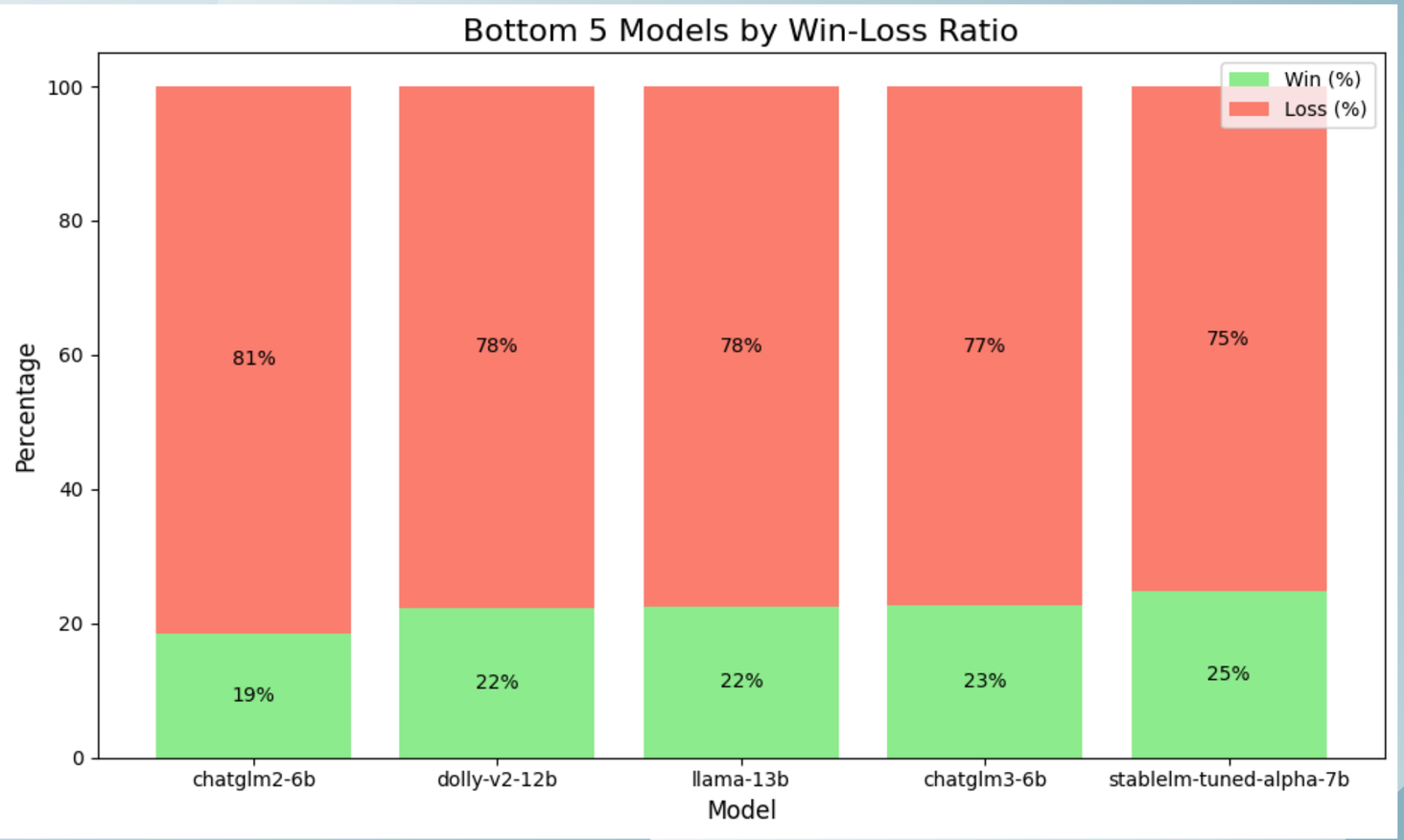
EDA



模型win-loss比率最高的前五名：

1. gpt-4-1106-preview
2. gpt-4-0125-preview
3. gpt-3.5-turbo-0314
4. gpt-4-0314
5. claude-1

EDA

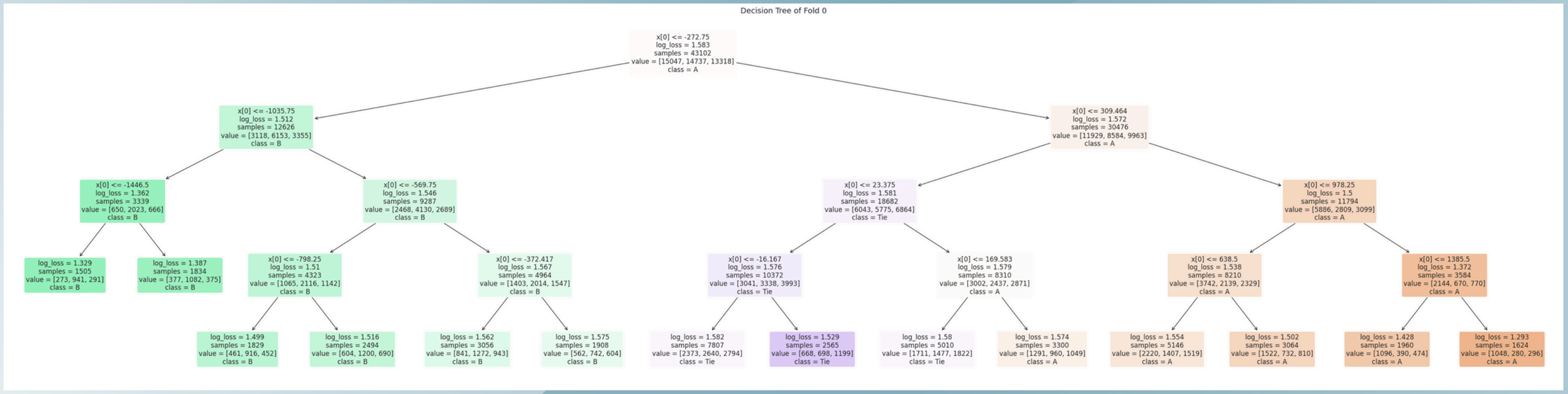


模型win-loss比率最低的前五名：

1. chatglm2-6b
2. dolly-v2-12b
3. llama-13b
4. chatglm3-6b
5. stablelm-tuned-alpha-7b

參考研究(1)-Decision Tree

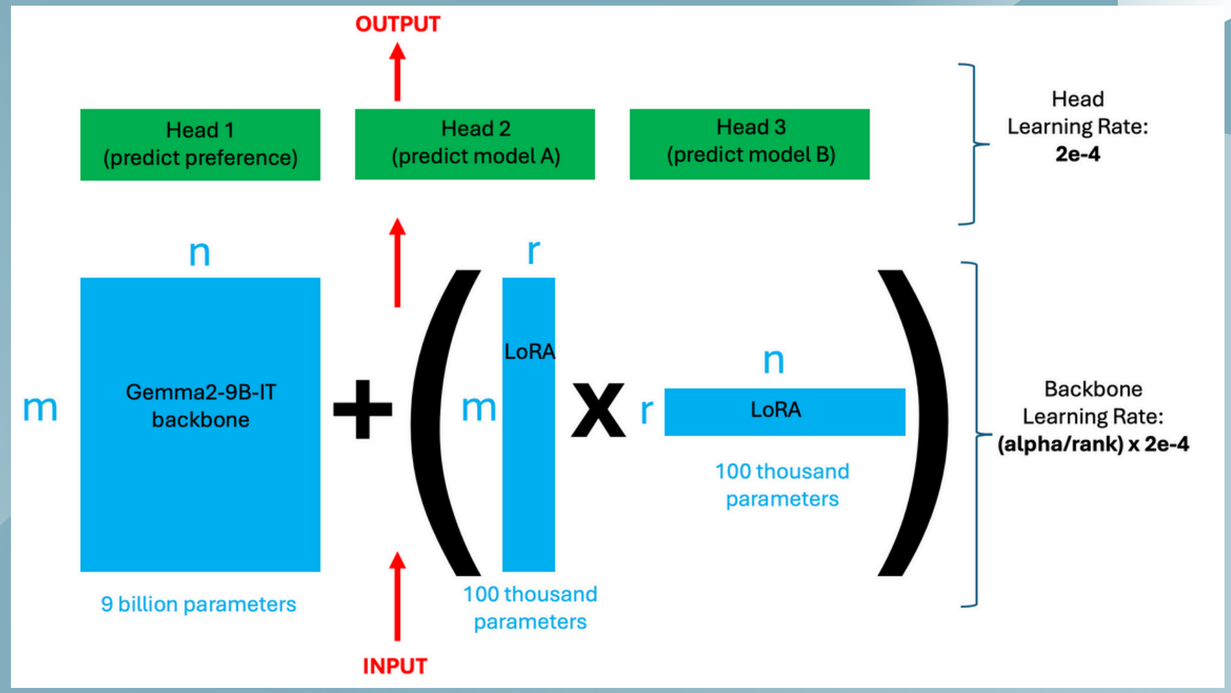
- Difference bucket mean prediction with 250 quantiles
- Log loss of 1.0511



https://www.kaggle.com/code/abaojiang/lmsys-detailed-eda#4.-ML-Baselines

參考研究(2)–LLM fine-tune Gemma2-9B-IT

- Backbone : Gemma2-9B-IT
- LoRA/QLoRA fine-tune
- Three head sperately predict
- Leaderboard : Gold medal 16th place, Private score : 0.98532



https://www.kaggle.com/competitions/lmsys-chatbot-arena/discussion/527596

參考研究(3)-Combine Gemma-2 9b & Llama-3 8b

- Backbone : LlaMa3 8b & Gemma2-9B-IT
- Seperately train both LlaMa3 8b & Gemma2-9B-IT
- Lora fine-tune
- Combine two models results : (gemma_results + llama_results) / 2
- Leaderboard : Silver medal, Private score : 1.01955

+

https://www.kaggle.com/code/jaejohn/lmsys-combine-gemma-2-9b-llama-3-8b

參考研究(4)-Gemma2 9b + Post Processing

- Backbone : Gemma2-9B
- Lora fine-tune
- Qunatization
- Post Processing : 如果 predict 的值超過某個 threshhold 就乘上 multiplier
- Leaderboard : Silver medal 21th place, Private score : 0.98614

```python
if POSTPROCESS:
    print("Postprocess applied ...")
    def adjust_probabilities(row, thresholds, multipliers):
        for col in thresholds.keys():
            if row[col] > thresholds[col]:
                row[col] *= multipliers[col]
        return row

    thresholds_075 = {
        'winner_model_a': 0.75,
        'winner_model_b': 0.75,
        'winner_tie': 0.75
    }
    multipliers_075 = {
        'winner_model_a': 1.125,
        'winner_model_b': 1.125,
        'winner_tie': 1.325
    }
```

https://www.kaggle.com/code/kishanvavdara/21st-place-human-bias-2-x-gemma-9b-4096-len

研究方法

1. 數據準備與差異化處理
   - 使用 Difference bucket mean prediction with 250 quantiles方法進行特徵分割。
   - 訓練模型學習不同bucket之間的分佈差異，提升模型對異質數據的泛化能力。
2. 輕量化模型訓練
   - 因訓練資源有限，將使用兩個較小型的backbone model (例如: Gemma2-2B 和 LLaMA-2 3B)，並分別進行 LoRA 或 QLoRA微調。
   - 使用 3 個seperate head分別訓練，提升專項預測能力。
3. 模型融合
   - 將 Gemma2-2B 和 LLaMA-2 3B 的輸出結果進行融合。
   - 融合結果能有效綜合兩個模型的優勢，減少單一模型的偏誤。
4. 後處理與校正
   - 設定閾值和 multiplier，進一步增強高置信度預測。

研究方法

方法	性能提升點	資源需求
參考研究 (1)	Bucket 處理提升泛化能力	中等
參考研究 (2)	單模型微調，準確率提升明顯	中等
參考研究 (3)	大型模型集成，顯著提升準確率	高
參考研究 (4)	後處理進一步提升模型表現	低
我們的研究方法	輕量化訓練 + 差異學習 + 集成模型 + 後處理，平衡準確率與資源	中低

優勢

- 資源友好：使用較小型的 Gemma2-2B 和 LLaMA-2 3B，避免高顯存需求。
- 準確率提升：結合多模型融合與差異學習，有效提升泛化能力。
- 高效後處理：透過簡單的閾值調整和加權校正，在推理後進一步優化結果。
- 易於擴展：未來資源充裕時，可替換更大 backbone 或擴展更多模型參與集成。

Thanks for Listening