

Applied Deep Learning (2022 Spring)

Homework 3

資管四 B07705015 劉鎮霆

Q1: Model

a. Model

- Model architecture

使用 huggingface 所提供的 google/mt5-small 作為預訓練模型¹，該多語言模型修改自 T5 架構並以 mC4 資料集進行預訓練。T5 採用 Encoder-Decoder 架構，encoder 端學習文句的特徵，decoder 端則類似 language model。

- How it works on text summarization

預訓練 T5 模型時，會將被挖空的句子輸入 encoder，並訓練 decoder 輸出被挖空的片段。當訓練或使用模型進行摘要時，則會將原文加上前綴 (如「summarize:」) 輸入 encoder，接著 decoder 端則以類似 language model 的方式逐一輸出摘要的 token。

b. Preprocessing

- Tokenization and data cleaning

首先將原文加上「summarize:」前綴，接著使用 transformers 套件中的 AutoEncoder 與 google/mt5-small 預訓練模型對原文與摘要 (僅 training dataset 與 validation dataset 包含摘要) 進行 tokenization，同時分別裁剪原文與摘要的長度 (此處設定原文最大長度為 256，摘要最大長度為 64)。

Q2: Training

a. Hyperparameter

由於我所使用的 GPU memory 不夠大，因此我將原文最大長度設為 256，摘要則設為 64。一開始使用 $5e-5$ 的 learning rate 時，我發現 ROUGE score 下降的非常緩慢，因此在其他設定相同的情況下逐漸增加 learning rate 並進行測試，之後發現超過 $5e-4$ 後繼續增加 learning rate (如 $1e-4$) ROUGE score 開始下降，因此將 learning rate 設為 $5e-4$ 。Batch size 則是在固定 learning rate 與其他參數後分別測試了 4、8、16 三種設定，發現設為 8 時表現最好。Hyperparameter 設定如下：

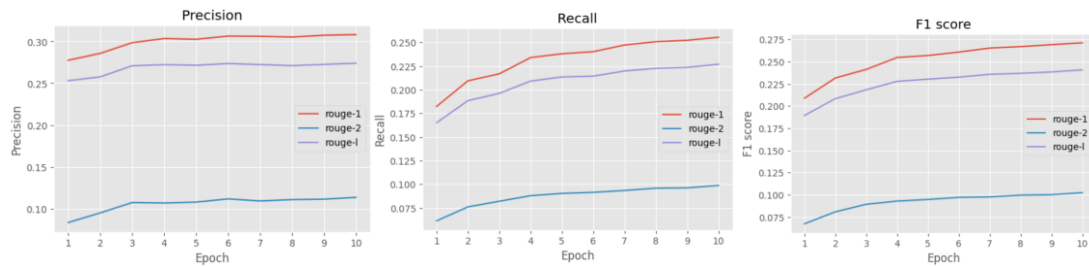
Pre-trained model	Max source length	Max target length
google/mt5-small	256	64
Optimizer	Learning rate	Weight decay

¹ <https://huggingface.co/google/mt5-small>

AdamW	5e-4	0
Batch size	# of Epochs	Scheduler
8	10	Linear decay

b. Learning curves

下圖為使用 public testing dataset (public.jsonl) 與 greedy generation strategy 畫出的 learning curve，可以看出隨著 epoch 增加，precision、recall 與 F1 score 皆有所增加。



Q3: Generation Strategies

a. Strategies

- Greedy：每次選擇下個字時，都選當前機率分布中有最大值的字
- Beam search：記錄最有可能的 k 個句子，接著每次再從目前的句子能夠和下個字連接出的句子中，選擇 k 個最有可能的句子
- Top-k sampling：從當前機率分布最大的 k 個字中隨機選一個字
- Top-p sampling：從機率總和大於 p 的最小集合中隨機選一個字
- Temperature：softmax temperature 並非 decoding algorithm，而是一種調整生成句子多樣性 (diversity) 的方式。將 $\text{softmax } P(w_t) = \frac{e^{s_w}}{\sum_{w' \in V} e^{s_{w'}}}$ 改為 $P(w_t) = \frac{e^{s_w/\tau}}{\sum_{w' \in V} e^{s_{w'}/\tau}}$ ， τ 是 temperature hyperparameter，值越大機率分布越平均，可以增加 generation strategies (sampling) 產生句子的多樣性，越小機率分布越集中，多樣性越低

b. Hyperparameters

- Try at least 2 settings of each strategies and compare the result

以下使用 public.jsonl，並以 F1 score 進行比較

- Greedy

	Rouge-1	Rouge-2	Rouge-l
Pure Sampling	0.191	0.058	0.167
Greedy	0.271	0.102	0.241

由上表可以看出 greedy 相比 pure sampling 有較佳的表現。

- Beam search

# of beams	Rouge-1	Rouge-2	Rouge-l
1 (greedy)	0.271	0.102	0.241
2	0.283	0.111	0.251
5	0.287	0.117	0.255

由上表可以看出 beams 數量越多，生成句子的 ROUGE score 也越高。

- Top-k sampling

k	Rouge-1	Rouge-2	Rouge-l
1 (greedy)	0.271	0.102	0.241
2	0.265	0.095	0.232
5	0.256	0.089	0.224
10	0.243	0.081	0.212
50	0.220	0.069	0.191

由上表可以看出本實驗中增加 k 的大小 (增加多樣性) 會得到較差的表現。

- Top-p sampling

p	Rouge-1	Rouge-2	Rouge-l
0.2	0.266	0.099	0.236
0.5	0.247	0.087	0.217
1.0 (pure sampling)	0.191	0.058	0.167

類似於 top-k sampling，增加 p 的大小也會得到較差的表現。

- Temperature

- Top-k sampling (k=5)

τ	Rouge-1	Rouge-2	Rouge-l
0.2	0.270	0.101	0.240
0.5	0.267	0.098	0.236
1 (original softmax)	0.256	0.089	0.224

由上表可以看出 temperature 越小 (多樣性低)，表現越好。

- Top-p sampling (p=0.2)

τ	Rouge-1	Rouge-2	Rouge-l
0.2	0.271	0.102	0.241
0.5	0.271	0.102	0.240
1 (original softmax)	0.266	0.099	0.236

類似於 top-k sampling，temperature 越小表現越好。

- Final generation strategy

在本實驗中，增加 sampling 的多樣性反而會得到較差的表現，其中 beam search (# of beams: 5) 的表現最好，因此最後以此作為 generation strategy。

Bonus: Applied RL on Summarization

a. Algorithm

- RL algorithm and reward function

使用 policy gradient fine-tune Q2 訓練好的 supervised-learning 模型，將 θ 表示為當前的模型參數、 θ' 為更新後的模型參數、 η 為 learning rate、 $\mathcal{R}(\theta)$ 為 reward function，使用以下方法更新模型參數：

$$\theta' \leftarrow \theta + \eta \nabla \mathcal{R}(\theta)$$

其中，將 N 表示為 batch size、 T_n 為生成句子的長度、 $R(\tau_{\text{sample}}^n)$ 為使用 sampling generation strategy (此處使用 top-p sampling) 產生第 n 個句子得到的 score、 $R(\tau_{\text{greedy}}^n)$ 為使用 greedy generation strategy 產生第 n 個句子得到的 score (此處作為 baseline)、 a_t^n 為 sampling generation strategy 於第 n 個句子產生的第 t 個 token、 s_t^n 為第 n 個句子在產生第 t 個 token 時模型的輸出、 $p(\cdot|\cdot)$ 為條件機率、 $\nabla \mathcal{R}(\theta)$ 可表示為：

$$\nabla \mathcal{R}(\theta) = \frac{1}{N} \sum_{n=1}^N \sum_{t=1}^{T_n} \left(R(\tau_{\text{sample}}^n) - R(\tau_{\text{greedy}}^n) \right) \nabla \log p(a_t^n | s_t^n, \theta)$$

其中， $R(\tau)$ 可表示為：

$$R(\tau) = \frac{1}{3} \left(\frac{\text{ROUGE-1}(\tau)}{0.271} + \frac{\text{ROUGE-2}(\tau)}{0.102} + \frac{\text{ROUGE-L}(\tau)}{0.241} \right)$$

由於對同個句子而言不同 rouge score 的大小不同，為了使不同 rouge score 對於 fine-tune 的影響相同，將此處的 rouge score 分別除以 Q3 中 greedy generation strategy 得到的 rouge score，並計算其平均。

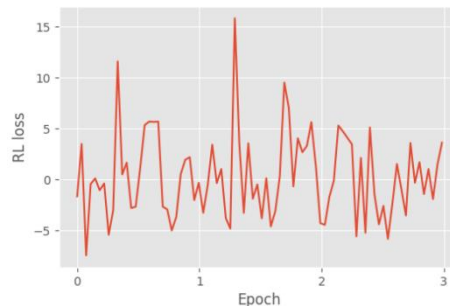
- Hyperparameters

Pre-trained model	Max source length	Max target length
model trained in Q2	256	64
Optimizer	Learning rate	Weight decay
AdamW	1e-5	0
Batch size	# of Epochs	Scheduler
8	3	Linear decay
p (in top-p sampling)	temperature	
0.2	1.0 (original softmax)	

b. Compare to Supervised Learning

• Learning Curve

由下圖可以觀察到每個 batch 的 RL loss 變化較大，沒有明顯的下降趨勢，這可能是因為產生文字時使用的是 sampling generation strategy，導致模型在訓練過程中仍然可能生成 reward 較差的句子。



• ROUGE score

使用 greedy algorithm 作為 generation strategy，比較 Q2 的 supervised-learning 模型與使用 reinforcement learning fine-tune 後的差異（使用 public.jsonl，並以 F1 score 進行比較）

	Rouge-1	Rouge-2	Rouge-l
Supervised learning	0.271	0.102	0.241
Reinforcement learning (fine-tuned)	0.275	0.108	0.245

fine-tune 後，模型在三種 ROUGE score 的表現都有所提升。

• Output texts

Ground-truth Title	Supervised Learning	Reinforcement Learning
陸男持刀挾持中學生 警員下跪脫衣要求替換人質	雲南挾持人質遭劫匪擊斃 疑是警察當時被狙擊手擊斃	雲南男持刀挾持學生 遭警狙擊手擊斃
科技股率美股下殺 美債殖利率飆升為風險資產敲響警鐘	科技股重摔 3.5% 科技股重摔 3.5%	科技股重摔 3.5% 道瓊大跌 559.85 點
義守大學前瞻未來 向全球學生招手	義守大學推動 AI 智慧科技 推動數位化學習環境	義守大推動跨域創新 推動跨域創新
免費送香檳氣泡酒！五星「台北萬豪酒店」3 餐廳搶推情人節套餐，威靈頓牛排必吃	情人節限定！五星「萬豪酒店」首推「偽出國」情人節套餐優惠	浪漫情人節！五星「萬豪酒店」首推情人節套餐優惠，加碼男男女女女女女女女必衝
讓老鹽埕變成生活聚場 邱承漢為新舊世代牽紅線	從鹽埕到鹽埕 邱承漢：從鹽埕到鹽埕	從鹽埕到鹽埕 從鹽埕到鹽埕的新故事
影／腦麻女臥床 50 年老父讓她「解脫」 北院籲請總統特赦	女兒牙痛吃止痛藥無效 法官籲總統特赦	老翁吃止痛藥無效 妻吃止痛藥無效無效

可以觀察到有些文章 fine-tune 後的模型能摘要得比較精確，如：

- 「陸男持刀挾持中學生...」：文章提到劫匪被警察擊斃，與 RL (reinforcement learning) 模型的摘要接近，而 SL (supervised learning)

模型的語意則較不通順，且變成警察被擊斃

- 「科技股率美股下殺...」中，SL 模型摘要出重複的句子，且較不通順

有些文章是原先的模型表現較佳，如：

- 「義守大學前瞻未來」：RL 模型摘要出重複的句子，且較不通順
- 「免費送香檳氣泡酒...」：RL 模型雖然前半段與標題接近，但後半段重複多次「男」、「女」，較不通順

也有些文章是兩者的表現都有待加強，如：

- 「讓老鹽埕變成生活聚場...」：兩者都生成重複的句子，較不通順
- 「影 / 腦麻女臥床 50 年...」：文章中提到老父由於身患腦麻的女兒吃止痛藥無效而將其悶死，但兩者摘要的語意皆和標題有所出入