

組長 GitHub ID : ming024

組員

- r08922080 資工所碩一 簡仲明
- r08921062 電機所碩一 黃健祐
- b04501127 土木四 凌于凱

1. Please give some examples predicted correctly and incorrectly respectively. At least one for each case is required. Screenshot recommended.

- **Context**

在歐洲，梵語的學術研究，由德國學者陸特和漢斯雷頓開創。後來威廉·瓊斯發現印歐語系，也要歸功於對梵語的研究。此外，梵語研究，也對西方文字學及歷史語言學的發展，貢獻不少。1786 年 2 月 2 日，亞洲協會在加爾各答舉行。會中，威廉·瓊斯發表了下面這段著名的言論：「梵語儘管非常古老，構造卻精妙絕倫：比希臘語還完美，比拉丁語還豐富，精緻之處同時勝過此兩者，但在動詞詞根和語法形式上，又跟此兩者無比相似，不可能是巧合的結果。這三種語言太相似了，使任何同時稽考三者的語文學家都不得不相信三者同出一源，出自一種可能已經消逝的語言。基於相似的原因，儘管缺少同樣有力的證據，我們可以推想哥德語和凱爾特語，雖然混入了迥然不同的語彙，也與梵語有著相同的起源；而古波斯語可能也是這一語系的子裔。」

- **Questions & Answers**

- 印歐語系因為哪一門語言而被發現？→ 梵語 (**correct**)
- 陸特和漢斯雷頓開創了哪一地區對梵語的學術研究？→ 德國 (**incorrect**)

2. Which hyperparameter(s) should be modified in order to reach better performance? Please explain why it works briefly.

The **learning_rate** and **warmup_steps** should be carefully selected. Without warm-up, the model could fall victim to early overfitting.

I modified **warmup_steps** from 512 to 16, and the mean F-score of public test set dropped from 0.81477 to 0.81116. Then, I adjusted **learning_rate** from 0.00005 to 0.0005. The score was 0.74318, showing the importance of learning rate in training.