Applying BERT on Facebook Children's Book Test
CS410

Team - Children Hooray
evanluo2 (Evan Luo), chiachi5 (Chia-Chi, Chen),
tiehchu2 (Tieh Chu)

Description

Our task is to apply BERT, a transformer-based machine learning technique for natural language processing, on Facebook Children's Book Test Dataset.

Facebook Children's Book Test (CBT)

In the Children's Book Test, a collection of children's books were gathered from the Project Gutenberg archives. Each question is constructed by taking 20 consecutive sentences from the book text, and leaving the 21st as the query statement. A word from the query is selected and masked, and the reader is tasked with selecting which word from the text (of the chosen type) should be used to fill this placeholder in the query.

` What is it ? '
answered he .
` The ogre is coming after us .
I saw him . '
` But where is he ?
I do n't see him . '
` Over there .
He only looks about as tall as a needle . '
Then they both began to run as fast as they could , while the ogre and his dog kept drawing always nearer .
A few more steps , and he would have been by their side , when Dschemila threw the darning needle behind her .
In a moment it became an iron mountain between them and their enemy .
` We will break it down , my dog and I , ' cried the ogre in a rage , and they dashed at the mountain till they had forced a path through , and came ever nearer and nearer .
` Cousin ! '
said Dschemila suddenly .
` What is it ? '
` The ogre is coming after us with his dog . '
` You go on in front then , ' answered he ; and they both ran on as fast as they could , while the ogre and the dog drew always nearer and nearer .
` They are close upon us ! '
cried the maiden , glancing behind , ` you must throw the pin . '
So Dschemil took the pin from his cloak and threw it behind him , and a dense thicket of thorns sprang up round them , which the ogre and his dog could not pass through . '

I will get through it somehow , if I burrow underground , ' cried he , and very soon he and the XXXXX were on the other side .
[gold]: dog
[answer canditates]: Cousin | cloak | dog | maiden | mountain | needle | path | pin | side | steps

Why this task

After taking the course, we find that there are unlimited possibilities to utilize machines to understand text content though it is still a challenging task. Therefore, we decided to apply the state-of-the-art BERT model to run on Facebook Children's Book Test Dataset in order to learn more about how it works while examining the best performance we can achieve.

Demo

(1) How to install the software.

(2) How to use the software.

(3) An example to allow a grader to use the provided use case to test the software.

Thank you!

Team - Children Hooray