Project Report

Lao Zishan
zishan.lao@student.kuleuven.be
r0910556

1 Project Introduction

The goal of this experiment is to investigate whether BERT, a neural transformer model, has implicit syntactic abilities. BERT is trained to predict certain words in a sentence that have been replaced by a mask symbol, without any explicit syntactic information. However, it performs well on various natural language processing tasks. The experiment will test the model's ability to predict the correct verb form for subject-verb agreement in sentences where the subject is not immediately adjacent to the verb. If the model can correctly predict the singular or plural form, it suggests that it has an implicit understanding of syntax. The experiment will use a small set of 50 sentences in English.

2 Description of Dataset

The dataset for the experiment is self-built, which have a similar structure as the examples provided, with a hierarchical structure that requires an understanding of the actual subject for proper agreement. The first column is the complete sentence, which has an actual subject and an attractor noun. The second column is the masked sentence, where the main verb is masked for predicting whether it should be in singular form or plural form. The third column is the expected word, and the fourth column is the unexpected word.

3 Coding Process

```
#Install the transformer library
!pip install transformers

# Disabling gradient computations
import pandas as pd
import torch
torch.set_grad_enabled(False)

# Pre-loading the BERT model:
from transformers import BertTokenizer
from transformers import BertForMaskedLM

MODEL_NAME = 'bert-base-cased'

tokenizer = BertTokenizer.from_pretrained(MODEL_NAME)
model = BertForMaskedLM.from_pretrained(MODEL_NAME)

#read in the dataset file
df=pd.read_excel('sentences.xlsx')
df.head()
```

```

# Define a function for BERT to predict the word form
def compareProbabilities(row):
  instance = row['Masked_sentence']

  expected = row['Expected']
  unexpected = row['Unexpected']

  # convert words to indices
  inputs = tokenizer(instance, return_tensors='pt')

  # get model outputs
  outputs = model(**inputs)

  # determine position of the masked token
  mask_index = int((inputs['input_ids'][0] == tokenizer.mask_token_id).nonzero())

  # determine expected and unexpected indices
  expected_index = tokenizer.convert_tokens_to_ids(expected)
  unexpected_index = tokenizer.convert_tokens_to_ids(unexpected)

```

```
  # extract logits for expected and unexpected tokens
  expected_logit = outputs.logits[0,mask_index,expected_index]
  unexpected_logit = outputs.logits[0,mask_index,unexpected_index]

  # correct prediction returns 1, incorrect returns 0
  if expected_logit>unexpected_logit:
    print('CORRECT', instance, '| prediction:',expected,'|',(expected, expected_logit),(unexp
    return 1
  elif unexpected_logit > expected_logit:
    print('WRONG', instance, '| prediction:',unexpected,'|',(expected, expected_logit),(unexp
    return 0
  else:
    #no clear preference
    print('WRONG NO PREFERENCE', instance, (expected, expected_logit),(unexpected,unexpected_
    return 0

# Iterate through the dataset using the function and compute the accuracy
correctCounter = 0
for index, row in df.iterrows():
  print(index)
  result = compareProbabilities(row)
  correctCounter+=result
print('Accuracy on dataset:', correctCounter / len(df))
```

4 Results and Analysis

It turns out that the accuracy on dataset is 0.94, which means that there are 3 wrong predictions out of 50 sentences.

The results of this experiment show that BERT is able to predict the correct subject-verb agreement with a high accuracy for English language, despite the presence of attractor nouns in the sentences. This suggests that the model has implicitly learned a certain notion of syntax, and is able to understand the hierarchical structure of sentences and determine the correct subject for subject-verb agreement.

However, it is worth noting that there are a few instances where the model's predictions are incorrect, such as in sentences 21, 24 and 47.

> 21 WRONG The pair of socks that my sister knitted for me [MASK] warm and cozy. | prediction: are | ('is', tensor(11.6084)) ('are', tensor(12.8458))

> 24 WRONG A bouquet of yellow roses [MASK] color and fragrance to the room. | prediction: lend | ('lends', tensor(-1.4092)) ('lend', tensor(6.9207))

> 47 WRONG The pair of shoes which she bought yesterday [MASK] missing. | prediction: were | ('was', tensor(14.5286)) ('were', tensor(15.1318))

In the first case, "The pair of socks that my sister knitted for me [MASK] warm and cozy", the model predicted "are" instead of "is". This suggests that the model may not fully understand that "pair" is a singular noun, even though it refers to multiple items.

In the second case, "The pair of shoes which she bought yesterday [MASK] missing", the model predicted "were" instead of "was".

In the third case, "A bouquet of yellow roses [MASK] color and fragrance to the room", the model predicted "lend" instead of "lends".

These three wrong cases suggest that the BERT model may have difficulty understanding collective nouns, which is a name for a group of people or things such as "family," "class," "pack," "bouquet," "pair," and "flock." Collective nouns usually take a singular verb, because they are singular in construction, but BERT recognized these collective nouns as plural forms. Other than these collective nouns, Bert seems to have excellent prediction of correct verb form for subject-verb agreement in sentences.

But it is important to note that this experiment was done on a small-scale dataset, and it would be beneficial to replicate this experiment on a larger set of sentences to further support these findings. Additionally, this experiment only focused on subject-verb agreement, future research could look into other syntactic phenomena to see how well BERT can handle them.

5 Conclusion

In conclusion, the results of the experiment showed that BERT is able to predict the correct subject-verb agreement with a high accuracy, despite the presence of attractor nouns in the sentences.

This suggests that the model has implicitly learned a certain notion of syntax, and is able to understand the hierarchical structure of sentences and determine the correct subject for subject-verb agreement, which is a significant finding as it highlights the ability of neural network architectures to learn syntax without the need for explicit syntactic structure in the input.

Furthermore, the ability of BERT to handle attractor nouns, which are known to be difficult for traditional models to deal with, adds to the strength of the model in understanding the complexities of natural language, which has important implications for NLP tasks such as question answering, machine translation, and text summarization. BERT's ability to understand syntax can improve the performance of these tasks, making them more accurate and efficient.

Overall, the results of the experiment demonstrate the potential of BERT as a powerful tool for natural language processing and the ability of neural networks to learn syntactic structure in a more implicit way.