CONTRADICTORY, MY DEAR WATSON

DATA 3402 Project

Joey Hussain

OVERVIEW

- The task, create an NLI model that assigns labels of entailment, neutral, and contradiction to pairs of premises and hypotheses in a multitude of languages as well as explore the use and applications of TPUs.
- The approach I took in this project was to use and compare three different models and make a comparison between them.
- The overall best model that I found was the xlm-roberta-large-xnli scoring almost 1.5x better than the other models.

DATA

- Type:
  - Input: Testing CSV File: id → premise → hypothesis→ lang_ab→ language
  - Input: Training CSV File: id → premise → hypothesis → lang_ab → language → label
- Size:
  - 4.02 MB
- Instances:
  - 12120 unique ids

PROBLEM FORMULATION

- Input / Output - I receive an ID number, a premise, and a hypothesis and I must determine the relation to the premise and hypothesis.
- Models used
  - xlm-roberta-large-xnli
  - bert-base-multilingual-uncased
  - bert-base-multilungual-cased

TRAINING

- I trained on Kaggle Notebook because I needed access to TPUs since the basis of the challenge wanted us to use it.
- Each model took 10 Epochs and it also used relu dense layers
- I stopped my model by using tf.keras.callbacks.EarlyStopping

PERFORMANCE COMPARISON

- xlm-roberta-large-xnli - Scored 0.92762
- bert-base-multilingual-uncased - Scored 0.62906
- bert-base-multilungual-cased - Scored 0.65158
- The Best Performing Notebook for this Kaggle challenges is 0.9767.
- Show/compare results in one table.

CONCLUSIONS

- Based on this project I realized that RoBerta pre-trained models are better optimize in dealing with semantic similarities over just the base Bert models.
- However, while my score was high with the RoBerta model there are ways to even increase that score further as I am still off of the highest scoring notebook on Kaggle

FUTURE WORK

- Now that I am familiar with NLI and using TPUs, I can further look into NLP and how you can find semantic similarities to train a machine to read, hear and then interpret text to give a thoughtful conversation in multiple languages possibly.