**Abstract**

**Introduction**

- About COPA

- Motivation of COPA (why is it important)

**COPA**

- The COPA problem

- COPA Data with examples

**BERT**

- The state-of-the-art results use large pre-trained language models like BERT, GPT, and RoBERTa.

- BERT is an open source project created by google.

- Google made two versions open source BERT\_base and BERT\_large (explain BERT base and large)

- It uses transformer and attentions heads

- Diagrams of BERT Transformers and encoders-decoders

**Our Model**

- Evolution of our Model (how did we improve our model)

* COPA question transformation. First model created two separate features that we not related.
* 2nd and 3rd models used a linear transformation to merge the two COPA statements into a single feature feed into BERT.
* Pre-train on MNLI only and then fine-tune on COPA

- Creating new data from MLNI

- Data Pre-processing (BERT tokenizer, [CLS], [SEP]., MNLI to COPA converter)

- Flow Chart of our model (Linear Transform Layer, BERT Layer, Customer Classifier with SOFTMAX)

- Paragraph about each layer in the model

**Future Work**

- Expand on the idea of using in-memory models and WordNET in our data pre-processing

- Can we transform "HARD" COPA statements into "Easy" statements using data pre-processing

- Use RoBERTa

**Results**

- Graph on model's accuracy compared to state-of-the-art results

- Graph of our model's loss and accuracy

- How we measure our model’s accuracy ( Avg of 10 consecutive experiments with the same hyper-parameters)

**Conclusion**

-Summary of our project and any final observations.