Summer Internship 2024

<<Interpret-CXR >>

Team members:
Faculty-in-charge: Dr. Sowmya Kamath S.

Department of Information Technology
National Institute of Technology Karnataka, Surathkal

May 29, 2024

- FineTuned GPT2 with our dataset
- findings_corpus – has list of tokenized sequences, all NaN values removed
- GPT2 Dataset Class takes tokenized list as input and ensures each sequence is truncated/padded also appends token_id at end of sequence
- collate_batch – member function of class used to pad
- Load Pretrained GPT2 model with a language modelling head for text generation

Takes 2 random sample reports from test split of the dataset and generates findings for it based on first 5 words of the model.

Generated Illogical Outputs

Used AdamW Optimizer to train the model for 3 epochs

Higher epoch number was leading to overfitting

Finetuning the model did improve its ability to churn out logical reports significantly

Saved the model and repeated the same for impressions corpus

NEXT

Setup a joint model

Integrating a **Swin Transformer** as a vision encoder into the pipeline that includes a language model. Our goal is to map the image features extracted by the Swin Transformer into the same embedding space used by the language model.

Input Image -> [Swin Transformer] -> 768-dimensional feature vector

768-dimensional feature vector -> [Linear Layer] -> Projected 768-dimensional vector

Projected 768-dimensional vector -> [Tanh Activation] -> Final 768-dimensional embedding

Final 768-dimensional embedding -> [Language Model] -> Downstream tasks (e.g., text generation, classification)

THANK YOU