

Project Proposal for Group 16

Choice of model

One of the most popular models in NLP these days is the BERT model. It has shown groundbreaking results in many classic tasks in language comprehension. Since transformers are part of the class curriculum, we chose to use an implementation of BERT for our project.

Unlike RNN, transformers rely purely on attention mechanisms, which considerably reduce computational time. That being said, we still need a framework that enables training on a GPU. We will probably use Pytorch for that purpose.

Choice of task

We propose to use one of the implementations of BERT provided by HuggingFace as detailed in [1] and apply it to some task such as sentence prediction, sentence classification, question answering or answering multiple choice questions.

At the time of writing of this proposal, we have not yet settled on any specific task but we intend to focus in one of the tasks listed above. This would be done to a general domain with the pre-trained weights provided.

Improvements

We want then to apply the chosen task to a different domain. For example, if we decide on sentence completion, we would then like to see how the model performs in completing sentences for a more specific domain like a medical domain.

For this, we would have to use transfer learning and fine tune the model by incorporating a medical corpus. After retraining this extended model, we would then like to compare our results in the new domain before and after fine tuning the model.

References

- [1] Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, R'emi Louf, Morgan Funtowicz, and Jamie Brew. Huggingface's transformers: State-of-the-art natural language processing. *ArXiv*, abs/1910.03771, 2019.