CS 584 Machine learning project

Feed-Forward vs LSTM vs Transformer models

Team members: Tristan Leduc, Paul Legout

Type of subject : Application

Problem's description :

RNN are commonly used to analyze sequential data. Analyzing this type of data requires consideration of the order of elements. To that extent, RNN have the ability to memorize small chunks of data and use it to predict the next item in line or to classify time series signals such as electrocardiograms in the medical field. Not only that, but this memory also makes this kind of network suitable for the analysis of language (NLP), in which the analysis of sequences provides information about semantic and context. LSTM are among the most used of this type, but they are prone to problems such as vanishing gradients, slow training and they struggle with long sequences. Transformers bring the concept of "Attention" to prevent this kind of problem by focusing on the relevant part. We will see the efficiency of this mechanism.

Process :

Implementation of a Feed-Forward, LSTM and Transformer neural networks and comparative analysis on several datasets covering multiple fields in order to show the versatility and the power of RNN, and particularly Transformer neural networks. Throughout this project, we will put transformers to the test, and compare their performance with LSTM and standard feed forward models.

Milestones :

Choice of 2-3 relevant datasets

<u>For each dataset:</u>

- Creation of a feed forward model architecture
- Creation of a LSTM model architecture
- Creation of a Transformer model

<u>For each model:</u>

- Training of the models
- Evaluation
- Performance comparative analysis

Conclusion on each RNN strengths and weaknesses. When RNN should be used, and which type of data is better suited for Transformer over LSTM.