Chapter 1

Objective of the project

The goal of this project is to understand how training a language model such as GPT2-Xl to convert it to COMET-DISTIL can affect its performance in other language modelling tasks when compared to its 'untrained' form. This will help us identify if this kind of training leads to a loss or a gain in capability for a language model. This is done by comparing the performance of the COMET-GPT2-XL[1] model against the 'naive' GPT2-XL on a variety of common modelling tasks and metrics.