

Study the Difference Between Published LLMs Based on Completing Truncated Text Pairs

Introduction:

Due to the recent popularity of Large Language Models (LLMs), there are many different LLMs published online, however, we are curious that how does each LLMs work and if they are affected based on their core theory or training data. So, we set up this experiment to explore the difference between these LLMs, we firstly find a public dataset Stanford Sentiment Treebank which contains the truncated text that could be our input to pass into LLMs to get a complete pair as outputs. Then we stored a new dataset that contains the original inputs, completed outputs from each LLM with labels. Then we built a classifier to check if we can identify which LLM produced a specific pair of text.