

Assignment 1

1. I want you to compare at least 10 human written texts (you can find from the internet Arxiv) with AI written text and see if
 - (i) the Zipflaw distribution exists in both.
 - (ii) Are there any differences observable based on the distributions of written text?

You can focus on one academic conference before 2023, take 10 papers such as NeurIPS, and compare with the 2024 version.

2. For your second assignment, I want you to extend this tool and let the user select what types of text to feed and compare their distributions.

You can feed them also into LLM and ask the differences between two types of text and see if LLM can identify any patterns.