

ReadMe

This data was used in the paper "Large Language Models as Partners in Student Essay Evaluation."

Folders #1, #2, #8, #11, #13, #14, #19, and #20 contain the evaluation results conducted using ChatGPT-4 for eight essays selected from a total of twenty-two. These were chosen based on four workshop themes, with two essays selected per theme. The eight essays include those cited in the qualitative analysis of Section 5 (Essays #1, #13, #19).

Due to the material size limit, only data for these eight essays are included. However, it is believed to be sufficient to fully understand the evaluation process.

Each folder is organized as follows:

1. Essay (File)
2. LLM Evaluation without Guidance (Folder): Results of six evaluations without any guidance.
3. LLM Evaluation with Rubrics (Folder): Results of six evaluations with specified rubrics.
4. LLM Evaluation by Comparison (Folder): Comparative evaluation results against the other seven essays.

All essays have been obtained with permission from the students who wrote them, under the condition of anonymity, and are only being publicly disclosed for research purposes.