

## I. Last week progress

- Looked into evaluation metrics for NLP
- Maintained the semantic similarity metric for evaluating final answers
- Added ROUGE for the evaluation of retrieved contexts
- Analyzed the resulting ROUGE scores

## II. Next week's to-do list

- Make the poster for the showcase
- Polish the interface and demo for the showcase if necessary
- Tune the parameters to improve the results if time allows