

Design and Complete a New Text Generation Task as a Researcher

(CSE 582; final project; last modified 4/12/2023)

Task: Design your own text generation task, collect your data, and select your preferred model to solve your task. Note: we only allow text generation tasks, no text classification tasks are allowed in this final project.

Period: [4/12/2023~5/5/2023](#)

Teams (flexible; team size = ``T’’):

- Individual researcher
- Group researchers (group member shares the same score)

What you need to submit (deadline 11:59pm on 5/5):

[URL of your github repository](#), including

- A pdf file describing your project
- Code for your project

Evaluation (max score 100; Points are deducted each time with a unit of 5 (for ``Individual researcher’’) or 5+T (for group researchers) for each of the following six dimensions):

- The novelty and soundness of your text generation **task** (Points: 0~15)
- How you collect **data**, scale, quality, data analysis (Points: 0~20)
- Why do you select that algorithm for your task; whether your implementation makes sense (Points: 0~15)
- If your **experiment** design and evaluation make sense (Points: 0~20)
- **Analyze** your experimental results, such as summarize some error patterns, why your model fails in certain cases, any other issues you observe in this task (Points: 0~20)
- The overall **writing** quality of your pdf file (Points: 0~10)

What you DO NOT need to do: (i) include baselines from other papers; (ii) use the memory-consuming, time-consuming, latest pretrained language models such as GPT3/4

You will receive: (i) your pdf with scores for each dimension and some comments/explanations; (ii) one or more best pdf reports in our class (anonymous)

Pls refer to TA, Shravya Chillamcherla (sjc6752@psu.edu), for more details on how to submit.