* Problem statement
  + Why use LLM?
  + Safe time, more diversity recipe
* Model
  + Mike email about gemma
  + Baseline
    - Construct comparable results
  + Fine-tine model
    - 统一 dataset
      * Total 10k (80% train, 20% eval)
      * Prompt
        + Get a more diverse prompt → overfit??
        + prompt: [Yitian Yan](mailto:yitianya@usc.edu) ✅

I have xxx, any cooking ideas?

How to cook xxx?

* + Auto evaluation
    - Add perplexity (research)

- fine-tuned-Flan-T5: 2.6157877445220947

- none-fine-tuned-Flan-T5: 3495272839643136.0

-GPT2.0-baseline: 7.774175799590266e+29

* + - Rouge
  + Human eval
    - Split out 3 dimension [Minhao Li](mailto:minhao@usc.edu)
    - 做一个三角形图 for presentation [Siyi He](mailto:siyih@usc.edu)
  + Generation sampling [Siyi He](mailto:siyih@usc.edu)[Minhao Li](mailto:minhao@usc.edu)[Yitian Yan](mailto:yitianya@usc.edu)✅
    - top k, top p
  + Masking [Siyi He](mailto:siyih@usc.edu)
  + Try out temperature

Paper

(看看去年的人都写了哪些部分)

abstract

Introduction (problem statement ish) [Yitian Yan](mailto:yitianya@usc.edu)

related work [Siyi He](mailto:siyih@usc.edu) ✅

Experiments:

* Gemma
* T5
* Bart
* Cross comparison/why gemma outperforms the other two (TBD)

conclusion/next step: [Minhao Li](mailto:minhao@usc.edu)

DDL 4.30

Bart:

* base
  + Train loss: 1.628245943069458
  + Perplexity: 2.091398000717163
  + eval\_loss: 0.7193598747253418,