CSE 538 2022 Project Ideas

Here is a listing of ten potential project ideas that you can choose from. I will keep adding more to this, so keep watching this space. You are also welcome to select your own project.

You need to identify a task, a dataset, a system that you will try to improve, the evaluation methodology you will use, and the analyses you will perform to understand the behavior of your model. You will also need to identify three key questions you will answer in doing your project.

The links below contain papers or github repo's which should give you a starting point for this. I will make you write a proposal that explicitly states a plan for your project.

1. Author sentiment analysis -- What does the author think about the main entity in the story?
https://github.com/StonyBrookNLP/PerSenT

2. Distributed Inference for large models -- What is the best way to set up inference over large language models on commodity hardware? For example, this paper shows how to run inference on the Bloom model using multiple (not commodity) hardware.
https://arxiv.org/pdf/2209.01188.pdf

3. Inferring story states -- Is this world state warranted in the given story?
https://arxiv.org/abs/2208.00329

4. TellMeWhy -- Why did the agent perform an action?
https://stonybrooknlp.github.io/tellmewhy/

5. Semantic Parsing for modeling formal specification
https://www.fsl.cs.stonybrook.edu/docs/mcfs/2022.lrec-1.233.pdf
https://github.com/StonyBrookNLP/specnfs

6. Natural Language Inference on Biomedical Data -- Generating sentences that explain the relation between two entities.

SuMe: https://github.com/StonyBrookNLP/sume

7. Human Language Modeling -- Characterizing the types of users a model fails on.
https://github.com/humanlab/HaRT

8. How many heads are necessary for different tasks? Do ablation studies to figure out how to prune models so they can be.
https://arxiv.org/abs/1905.10650

9. Emotional Reactions -- What emotional reactions did an event engender in a story?
https://github.com/StonyBrookNLP/emotion-label-semantics

10. Syntactic vs. Semantic Failures -- Do syntactic failure necessarily lead to semantic failures? Relate probe failures to task failures.

11. Controllable Generation -- Decoding with constraints. Can you extend this to richer types of constraints or show utility for some task?
https://arxiv.org/abs/2010.12884