

Project Timeline

Find a team! Find a project! (Today – October 5th)

Project Proposal – October 5th (15 %)

250-400 words

Project Presentation – November 16th – December 7th (20%)

Present whatever results you have

Project Report – December 7th (40%)

Completed projects

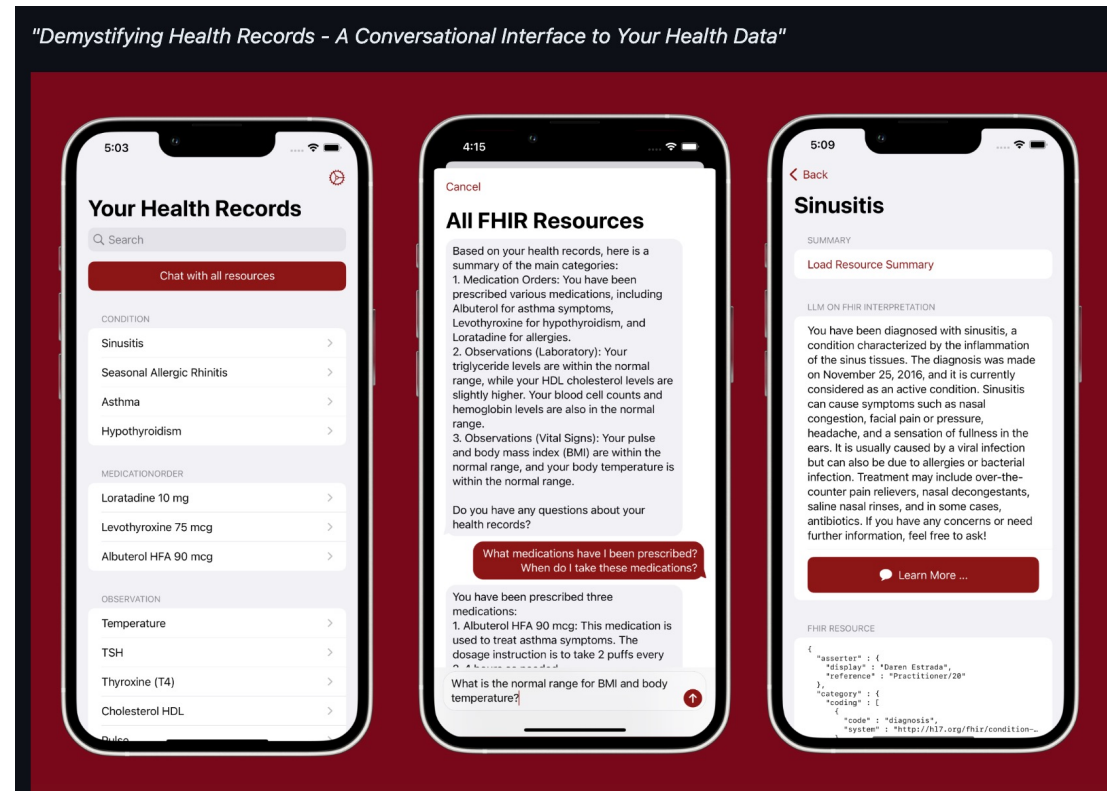
Getting started with project ideas

- **What are you interested? What are you currently researching?**
 - Is there something new you'd like to explore that's different from your current research
 - Is there a project that is inspired by your current research
 - Is there a new skill you'd like to gain
- **Finding overlap with your team members**
 - Many of you have done or do ML4H research, some may already have access to private datasets, varying levels of skill/experiences. Try to find a blend of interests within your group, or divide the labour based on the goals/expertise that each person has
 - Mix of clinical and technical people in the class!
- **Dataset/disease/modality ? Methods ? Analysis ? What came first?**
 - Think about what interests you most, then go from there:
 - "I want to work with unstructured text data (ie: EMR)" -> LLMs -> MIMIC CCU
 - "I'm interested in multimodal learning" -> LLMs + CXR -> MIMIC CXR
 - "I want to learn about time-series models" -> Transformers / ODE-RNNs -> HAR data
 - We can help you brainstorm during OH

Some Ideas

- FHIR via LLMs on MIMIC
 - Generating patient-friendly summaries of complex EMR data using LLMs
 - <https://github.com/StanfordBDHG/LLMonFHIR>
 - <https://physionet.org/content/mimic-iv-fhir-demo/2.0/>
- Blending publicly available LLMs for healthcare tasks
 - <https://ai.meta.com/blog/dinov2-facet-computer-vision-fairness-evaluation>
 - <https://ai.meta.com/blog/seamless-m4t/>
 - <https://ai.meta.com/llama/>
 - <https://github.com/facebookresearch/segment-anything>
- Evaluating biases in healthcare tasks using LLMs – comparing different LLMs to see which ones are best/worst
- **Past projects:** <https://csc2541hf-2021.github.io/>

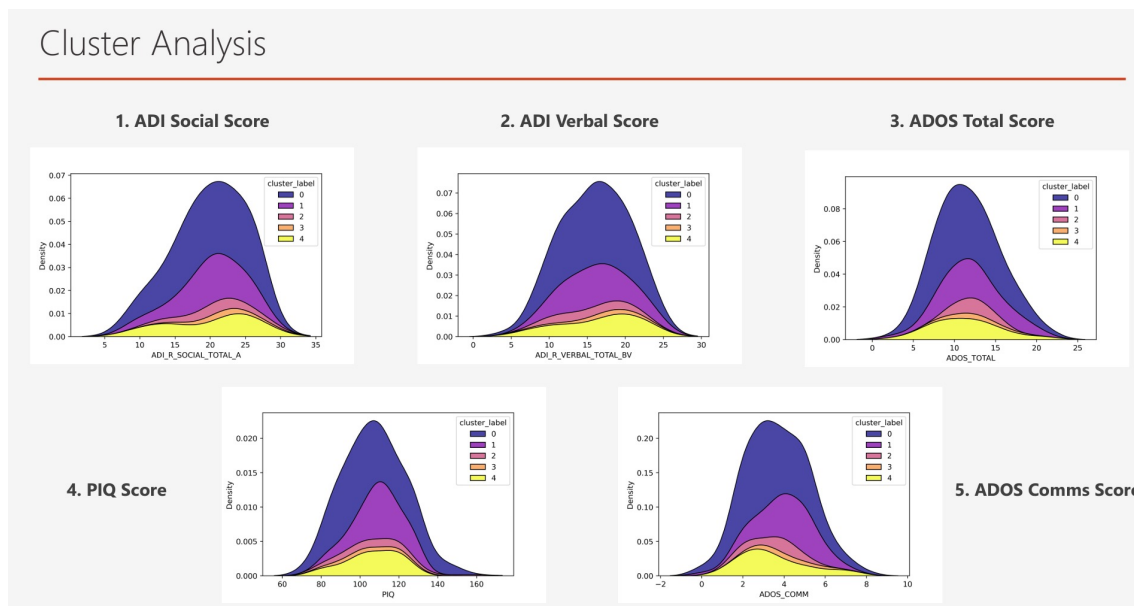
FHIR via LLMs on MIMIC



The application connects to the OpenAI GPT API to interpret FHIR resources using the GPT suite of large language models.

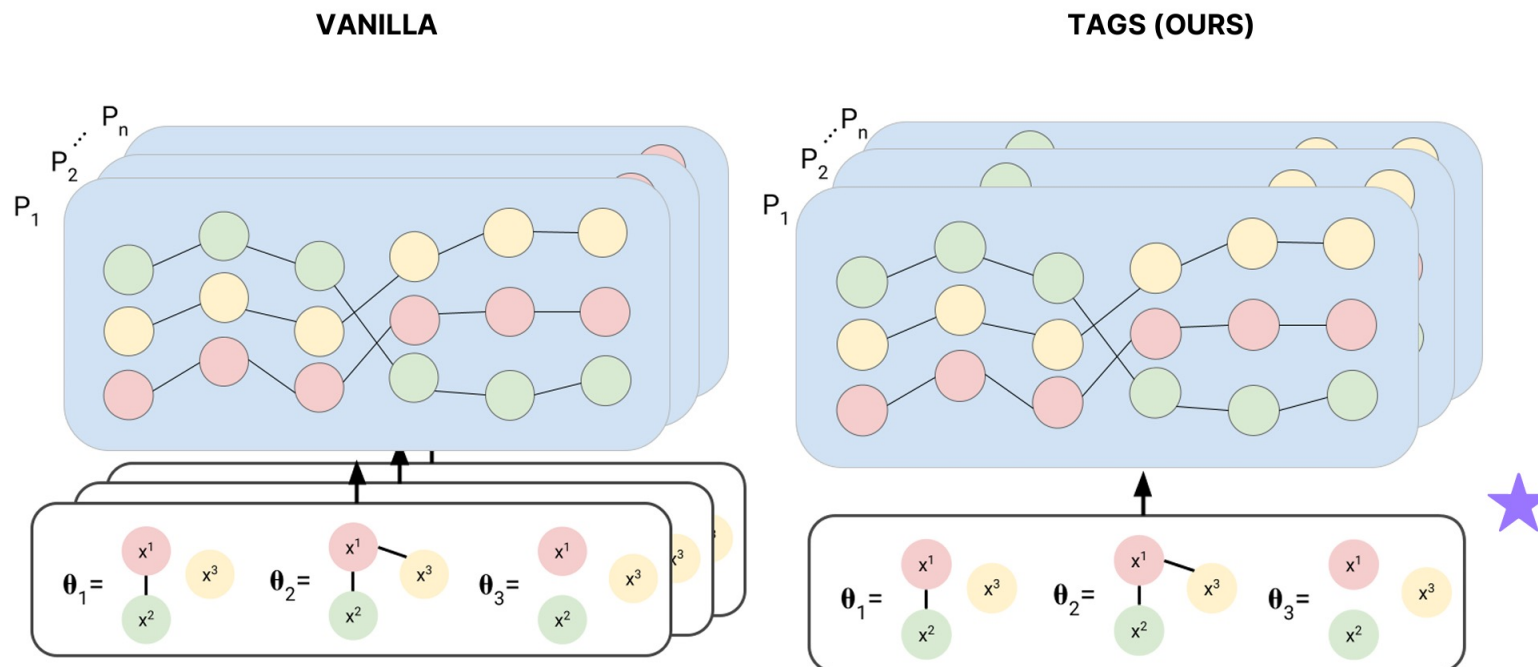
Multimodal ASD Subtyping (Banthia, Kolagati, Sun, and Deng ML4H 2021)

- Combining multiple modalities (IQ scores, ASD interview scores, FMRI) to identify novel subtypes of ASD
- Identifying similarity networks within each modality then combining them with similarity network fusion:



Hierarchical State Space Modelling (Nagaraj, Harrigan, Pokhrel, ML4H 2021)

- Unsupervised SSM to combine population-level and individual-level states (shared pool of states) to identify similarities between individuals



Project resources

- Cancer:
 - COMMPass study: <https://themmrf.org/finding-a-cure/our-work/the-mmrf-commpass-study/>
 - Starter code for processing data from the study: https://github.com/clinicalml/ml_mmrf
- Parkinson's disease
 - <https://www.ppmi-info.org/>
- **APPLY EARLY FOR ACCESS!**
- **Using ChatGPT (with caution) DEMO**
- **Find more at: <https://csc2541-2023.github.io/projectresources>**

Project Expectations

FAQ in Office Hours:

1. Is ____ a good project idea?

- a) Come see Rahul or Sujay in OH to get feedback on your projects

2. How novel should our project be?

- a) We aren't expecting major ground-breaking publications!
- b) Explore an interesting healthcare dataset, problem, method, or build your own

3. Can I re-implement existing methods?

- a) Yes and no. If you are re-implementing existing methods, we expect more analysis (ie: comparisons, identifying edge cases, pushing methods to their limits to learn more about what works well and when)
- b) If you come up with your own method, then we expect some existing methods to compare to

4. Are negative results OK?

- a) Absolutely!

Unstructured Time

Find groups if you have not done so already

Use this time to ask us questions about your projects / find datasets, etc...