Until ~7/20

Implementation

- Analysis scripts for codebook hierarchical structure (with visualization).
- Preprocessing scripts for annotation data:
    - Annotation statistics.
    - NLP structure alignment.
    - Visualization for statistics.
- Mapping script for inconsistencies between annotation and codebook labels.
- Data preprocessing script for training models.
- Text-to-graph conversion scripts:
    - Syntax dependency graph.
    - Semantic role graphs: weighted, predicate, clause-anchored variants.
- GCN implementation scripts.
- Error analysis scripts.
- Pre-class analysis scripts for each model type.

Documentation

- GitHub documentation for scripts (ongoing).
- Paper draft (ongoing; awaiting extension from analysis).

~7/28

Documentation & Investigation

- Lab connection meeting notes.
- Extended graph-based paper draft with two improvement strategies:
    1. **Method-side variations**:
        - Analyze which graph type benefits which intent classes.
        - Test context integration (adjacent only, whole message, symbolic propagation) (on hold).

        - Compare graph learning methods (GCN, GAT, etc.) (on hold).
        - Examine how symbolic context supports intent inference (ongoing).
    2. **Data/label optimization**:
        - Merge/remove consistently low-performing labels.
        - Test optimized datasets from team members.
        - Add multi-label support for overlapping intents.
        - Evaluate sentence vs. clause-level modeling trade-offs.

~8/4

Investigation & Implementation

- Reviewed optimized datasets from lab members:
    - Message-level annotations.
    - Combined label sets.
- Investigated public datasets with intent annotations:
    - DailyDialog.
    - CLINC150.
    - MedDialog (conversational, but lacking explicit intent labels).

Documentation

- Extended paper draft sections on:
    - Context-involved classification.
    - Alternative graph learning models.

Implementation

- Refactored scripts for cross-dataset reusability (consistent I/O structure).
- Facilitated easier downstream analysis and model generalization testing.

Other Documentation

- Drafted discussion section for goal paper.

~8/11 (Ongoing)

Documentation

- Identified ambiguous usage of “intent” in NLP and drafted discussion for graph-based paper.

Implementation

- Preprocessing scripts for:
    - Classification using optimized label hierarchies.
    - Data preprocessing for model training.

Investigation

- Reviewed prior studies on different intent types and their implications for symbolic modeling.