Overview of Tasks

Group Presentations

Jin Zhao

Brandeis University

March 2, 2025

Today's Session

Group task presentations continue

Format:

- 10-15 minutes per group
- 5 minutes Q&A / feedback
- Peer feedback forms

Goals:

- Share your annotation task design
- Get feedback from peers and instructors
- Learn from other groups' approaches

Presentation Content

Each group should cover:

1. **Task definition:** What are you annotating?
2. **Motivation:** Why is this task important/interesting?
3. **Data source:** What corpus are you using?
4. **Annotation schema:** What labels/tags?
5. **Example annotations:** Show 3+ annotated samples
6. **Anticipated challenges:** What do you expect to be hard?
7. **Tool choice:** What annotation tool will you use?

Feedback Guidelines

When giving feedback, consider:

- Is the task clearly defined?
- Are the categories well-motivated?
- Do the examples make sense?
- Are there potential ambiguities in the schema?
- Is the scope realistic for the semester?
- What edge cases might they encounter?

Be constructive:

- Point out potential issues
- Suggest alternatives
- Ask clarifying questions

Common Issues to Watch For

Schema design:

- Overlapping categories
- Missing important cases
- Too many fine-grained distinctions

Data selection:

- Corpus too small or too large
- Licensing issues
- Data quality concerns

Scope:

- Task too ambitious for timeline
- Task too simple for meaningful project
- Unclear success criteria

Upcoming Deadlines

Project milestones:

- **Week 9 (Mar 11):** Draft 1 guidelines + pilot annotations
- **Week 10 (Mar 18):** Draft 2 guidelines
- **Week 11 (Mar 25):** IAA evaluation
- **Week 13 (Apr 15):** Gold standard dataset

HW 2: Assigned this week (Data Wrangling with Pandas)

Let's begin!

Group presentations follow

Presentation Schedule

Today's presenters:

1. Group 1: [Topic TBD]
2. Group 2: [Topic TBD]
3. Group 3: [Topic TBD]
4. Group 4: [Topic TBD]

(Actual group names and topics to be filled in)

Discussion: Common Themes

After presentations, let's discuss:

- What annotation types are most common?
- What challenges do multiple groups share?
- How do different groups handle ambiguity?
- What tools are people using?
- What can groups learn from each other?

Lecture 13 (Mar 4): LLM-Based Annotation: Zero-shot & Few-shot

Topics:

- Zero-shot annotation with LLMs
- Few-shot in-context learning for annotation
- Prompt design patterns
- Model selection for annotation
- Structured output formats
- Cost estimation and optimization

Reading: Gilardi et al. (2023) – ChatGPT vs. crowd-workers

Questions?

Office Hours: Wednesdays 1-3pm, Volen 109

✉ jinzhao@brandeis.edu