Annota: Peer-based AI Hints Towards Learning Qualitative Coding at Scale

🙂 Dustin Palea, Giridhar Vadhul, David T. Lee

- **How might large classes be an advantage** rather than a disadvantage for experiential learning?
- We repurpose label aggregation algorithms **to provide peer-based hints that augment learning of complex work** rather than for micro-task work.
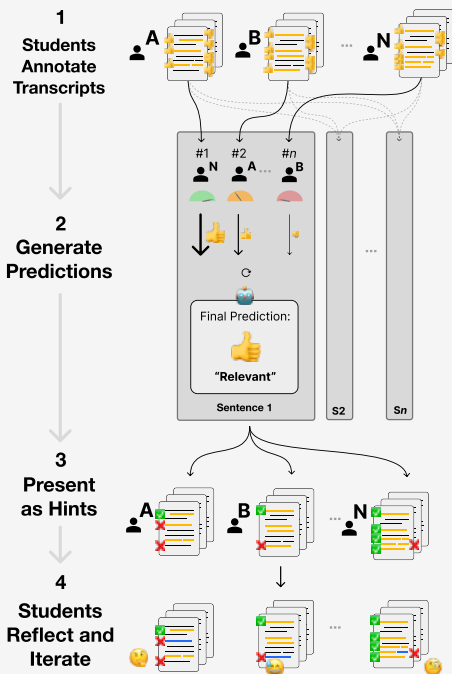
Intro & Context

- Learning qualitative analysis (QA) requires **personalized feedback** and **in-depth discussion**, but **educators cannot provide this in large classes**.

- To overcome this challenge, we introduce a **learnersourcing method** that builds on the **Dawid-Skene expectation maximization algorithm (DS-EM)** to generate **peer-based AI hints** that support students in one aspect of QA: **determining what sentences are relevant to the research question**.

- The **Annota** system was deployed in a **large business strategy class (N = 122)**. Students each analyzed **8 stakeholder transcripts** in relation to **research questions**, including: *Strengths & Weaknesses*, and *Opportunities & Threats.*

- We measure the quality of hints with a **Macro-F1** score relative to **expert labels/annotations**. We understand the **value to student learning** through **surveys**, **interviews**, and an **inductive qualitative analysis** process.

Peer-Based AI Hints in Annota
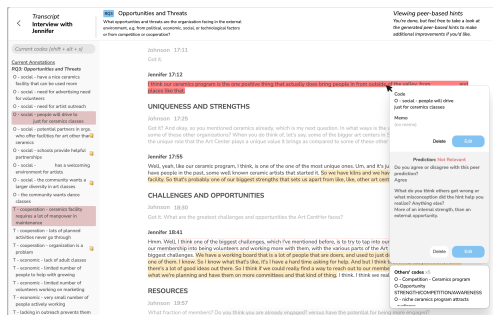
Can **collaborative generative AI** be used to facilitate **peer learning** through **peer-based AI hints?**

1 Students Annotate Transcripts

2 Generate Predictions

Final Prediction:
👍
"Relevant"

Sentence 1   S2   Sn

3 Present as Hints

4 Students Reflect and Iterate

Results

- Annota predictions are **similar to the best students** and **far better than the average student or majority vote** in Macro-F1 scores, with quality gains primarily coming from **strong positive recall and negative precision**.

- Accurate feedback **only requires ~15 annotators** and **assigning ~2 transcripts**.

- Peer-based AI hints **helped students improve** their **understanding of research questions**, more carefully **examine their transcript annotations**, and improve their understanding of when they were **over-annotating or under-annotating** the transcript.

Red ticks indicate the passing of an average length transcript.
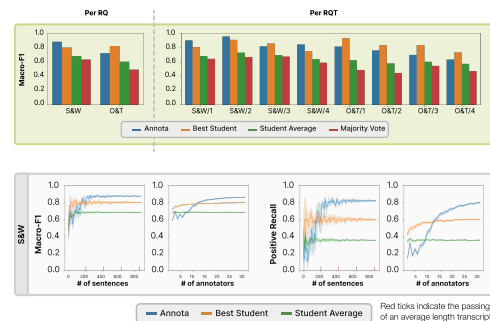
Collaborative Generative AI for Co-Creative Learning

- How might *collaborative* generative AI be used **for learning** the more **complex and interpretive** aspects of **qualitative analysis** (QA), beyond simply determining relevant sentences?

- After determining relevant sentences, they must be appropriately coded to capture their value relative to answering the research question. Could GenAI help students **better articulate their intent** behind making an annotation, for peers to more easily build upon their work in the later stages of QA?

- Discussion is an integral part of QA. Could GenAI **facilitate discussion across large numbers of students** to surface their unique perspectives and identify areas of disagreement for discussion?

🙂 Tech4Good Lab   UC SANTA CRUZ   BE Baskin Engineering