

CSE 447 Final Project Proposal

Team Members

- **Sam Shin** (NetID: *seunguk*), **Class Number:** CSE 447
- **Sebastian Liu** (NetID: *LL57*), **Class Number:** CSE 447

Project Information

- **Project Type:** Default 447
- **Task Name:** OpenBookQA
 - **Citation:**

Mihaylov, Todor, Peter Clark, Tushar Khot and Ashish Sabharwal. "Can a Suit of Armor Conduct Electricity? A New Dataset for Open Book Question Answering." Conference on Empirical Methods in Natural Language Processing (2018).

- **Benchmark:** https://leaderboard.allenai.org/open_book_qa/submissions/public
- **Why we pick this task:** We are fascinated by the concept of open book exams in the context of AI. Since open book exams are a classic way to test human understanding, assessing machines' ability to "understand" and apply core science facts correctly seems like a good way to measure their intelligence. We think this will be a challenging yet rewarding task to work on.
- **Research Plan:**
 - Get a deep understanding (week 1)
 - Read the paper understand its motivation and approach
 - Browse through the papers that cite this paper understand OpenBookQA's impact on the field
 - Survey around the leaderboard and get a list of popular solutions to OpenBookQA
 - Solutions Deep Dive (week 2)
 - Pick a few solutions that we think are interesting and try to understand them in depth
 - Analyze the pros and cons of each solution
 - Propose a list potential improvements to the solutions
 - Implementation & Experimentation (week 3 & 4)
 - Pick one solution we want to improve on
 - Implement our proposed improvements
 - List out the experiments we want to run
 - Run experiments and analyze the results