

Commonsense Question Answering with ARC

Prof. Dr. Steffen Staab
Dr. Corina Dima
M.Sc. Anastasiia Iurshina

First meeting: [to be defined]
Next meetings: [to be defined]



*Which property of a mineral can be determined just by looking at it?
(A) luster [correct] (B) mass (C) weight (D) hardness*

*Which technology was developed most recently?
(A) cellular telephone [correct] (B) television (C) refrigerator (D) airplane*

Target Group

This seminar targets students interested in Artificial Intelligence, Machine Learning and Natural Language Processing.

Topic

A challenge for current question answering systems are questions that require **reasoning** or **commonsense knowledge**. E.g., in the first question above the system needs to figure out that *luster* can be determined by looking at something, and not be fooled by the strong correlation between the words *mineral* and *hardness*.

The [AI2 Reasoning Challenge \(ARC\)](#) is a dataset containing a collection of 7787 natural language science questions with multiple choice answers. In this lab we will explore several methods for tackling the ARC challenge, ranging from IR and PMI-based methods to neural network models like LSTMs and transformers. You will be able to compare your solution to the ARC public leaderboard, see how your system measures up against the public submissions and - why not? - even show off by publicly submitting your results.

Procedure

Participants will report and discuss their progress weekly. At the end of the lab students will give their final presentations. At the beginning of the lab, the structure of the presentations and that of the reports will be discussed with the lecturers. The reports must be handed in before the final presentations and before the end of the summer term. The teaching language is English. All presentations and seminar reports will have to be delivered in English.