Project 1: Word Embeddings/Recurrent Neural Networks

This is a single-person project. It is fine to discuss your process with other students but make your own decisions and write your own code, run your own experiments, and create your own presentation.

Our task is common sense question answering with the CommonsenseQA dataset.

Architectures

We use two architectures:

1. Word embeddings (word2vec, GloVe, or fastText) together with a classifier (2-layer with ReLU non-linearity). Train only the classifier.
2. A 2-layer RNN architecture (LSTM or GRU, use the PyTorch implementations), and a two-layer classifier with a ReLU. For the input layer, use the same word embeddings as in architecture 1. Train the whole network end-to-end.

Deliverables

First, a Jupyter notebook with documentation, code, and a link to your Weights & Biases view or report. Second, the slides of your presentation as a PDF. Check the deadlines on Ilias. Late submissions result in grade deductions.

Data

Use the CommonsenseQA dataset from Hugging Face with the dataset splits as instructed in class.

Jupyter Notebook

Start every section with all your decisions and justifications. Include the following sections:

- Introduction
- Setup
- Preprocessing
- Model
- Training
- Evaluation
- Interpretation

For details, see the grading checklist. The notebook should be runnable and reproduce your results.

Presentation

5 minutes. Don't explain the task (everyone did the same task). Content (~1 slide per section):

- Preprocessing
- Input/output format
- Network architecture
- Experiments
- Results
- Interpretation of results