Assignment 4

Simon Clematide

Department of Computational Linguistics
University of Zurich

Machine Learning for NLP 1

Assignment 4: Paper Dissection

Identify an interesting and high-quality (short) NLP paper

- Interesting: Landmark paper mentioned in lecture/reading
- Interesting: SOTA paper listed in https://nlpprogress.com
- If paper is long and covers many Machine Learning approaches, focus on the best or simplest setup

Understand the paper

- Read the paper "quickly and efficiently"
- Go along the IMRaD schema (next slide)
- If you don't understand some concepts, try to find introductory resources (WP pages, quora, book chapters, blogs, videos) that help.
- But do not waste too much time into researching things that are totally unclear. Try to formulate/pinpoint what you don't understand and what is unclear.

IMRaD: Introduction, Methods, Results and Discussion[1]

Efficient reading order may not be linear order

- Abstract
- Conclusion
- Look at examples/figures/tables
- Introduction
- Methods
- Results
- Discussion

[1] https://francescolelli.info/thesis/read-scientific-papers-quickly-and-effectively/

Writing Your Paper Dissection: Max 2 Pages

- What is it about? What problem does it try to solve? Why is it interesting?
- Which ML methods are used? What is the main innovation of the paper?
- What are the take aways?
- What are possible problems of the approach? Think critically!
- You can add a mindmap if you like them
- You can also copy/paste the most important figure/table!
- What does one need to know for understanding the paper? Add the resources that were helpful for you.

Student Short Talks

- 8 minutes + 2 minutes questions
- In 3 slots in class, 3 slots in tutorial in November/December sessions
- Or: create a short screencast (e.g. with Screencastify) for "future" students (no perfectionism asked for!); e.g. a walkthrough to a code example

Topics

- A short paper: a technical or social/ethical aspect of ML in NLP
- A technical topic: GPU/TPUs; hierarchical Softmax; GloVe technical details; different optimizers (Adam); walkthrough of the code of a paper (Papers with Code)

Organization and Deadlines

Communicate your topics and suggestions via Feedback-Forum in OLAT

- Talks: Reply ASAP in forum thread "Student Talks" in OLAT and email me at the same time.
- Paper dissections: Thursday 19.1.2023 23:59: Hand-in your PDF in OLAT
- Screencasts: Thursday 19.1.2023 23:59: Hand-in Link to screencast in OLAT