



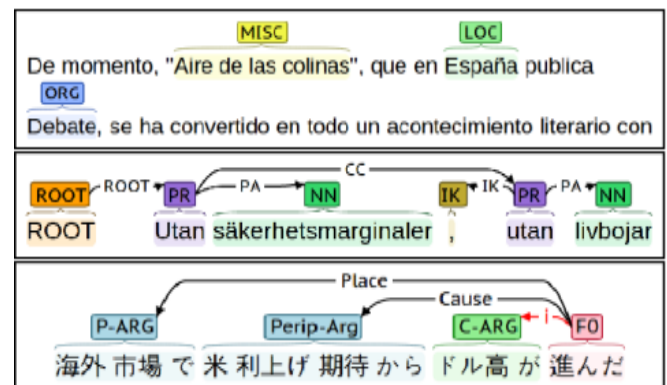
Online learning for interactive annotation

Motivation

With the rise of machine learning in Natural Language Processing, more and diverse annotated corpora are needed. Creating these is an expensive, time consuming and difficult undertaking. In order to speed up this process, we investigate interactive annotation in the open-source INCEpTION project (<http://tiny.cc/inception>). Our annotation editor offers machine learning based recommenders that suggest possible annotations to the user. Recommenders are retrained in the background and thereby improve their suggestion quality. A problem with this approach is the training time, as of right now, recommenders are often retrained from scratch. The goal of this thesis therefore is to research and evaluate *online learning algorithms*, i.e. algorithms that can be used to update models by only training on newly incoming data, reducing the need to retrain from scratch.

Task Description

- Investigate and apply state-of-the-art online learning algorithms (especially for deep learning)
- Evaluate them on a wide range of NLP tasks in a simulated interactive environment
- Integrate and evaluate promising candidates using INCEpTION, our annotation platform



References

- Losing V, Hammer B, Wersing H. Incremental on-line learning: A review and comparison of state of the art algorithms. *Neurocomputing*. 2018;275:1261-1274.
- Steven C.H. Hoi, Doyen Sahoo, Jing Lu, Peilin Zhao. *Online Learning: A Comprehensive Survey*
- Klie, J.-C., Bugert, M., Boullosa, B., Eckart de Castilho, R. and Gurevych, I. (2018): *The INCEpTION Platform: Machine-Assisted and Knowledge-Oriented Interactive Annotation.*

Contact

Analysis



Programming



Literature



Prof. Dr. Iryna Gurevych

Jan-Christoph Klie

thesis@ukp.informatik.tu-darmstadt.de