



Emoji for Natural Language Processing

It is not surprising that using glyph information in logographic languages helps to solve NLP tasks. However, the majority of natural languages are phonogram languages, which are represented in alphabet systems. NLP tasks in phonogram languages will benefit much less from the use of the glyph information.

Luckily, nowadays, a new global language appears: **emoji**. Emoji can be understood by anyone in the world no matter what the mother tongue is. Emoji can help people express their emotions more intuitively in communication. Beyond communication in daily life, some people are trying to use emoji in lifetime.

There is a mapping between natural language and emoji. It is intuitive that emoji should help us to solve NLP tasks in any language. In this thesis, we will study a wide range of NLP tasks, including tagging, sentence pair classification, single sentence classification tasks, dependency parsing, and semantic role labeling.

While we already have some ideas about how such algorithms might look, we are interested in your vision and ideas.

Requirements: Prior knowledge in machine learning, computer vision, or nature language process. We will have weekly meetings to address questions, discuss progress, and think about future ideas.

Interested? Please contact us for more details!

Contact

- Ye Wang: wangye@ethz.ch, ETZ G63
- Zhao Meng: zhmeng@ethz.ch, ETZ G61.4

