Title:
Cross-lingual transfer via prompt translation and multilingual language models

Contact:
jeremy.barnes@ehu.eus

Brief Description:

For a large number of classification tasks, methods that fine-tune large pretrained language models with large amounts of labelled data lead to state-of-the-art results. However, this requires 1) large amounts of data and 2) expensively fine-tuning these large models.

Prompt based learning has emerged as an alternative, where the main idea is to cast the target task into a formulation which allows us to exploit the knowledge implicit in a pre-trained LM, often using some label verbalization or question answering pattern.

For many under-resourced languages, however, we do not have large amounts of labelled data. However, given the multilingual ability of many modern LMs (multilingual BERT, XLM RoBERTa, mT5), it would be interesting to see if it is possible to automatically translate prompts to perform a task in an under-resourced language without having to create these large new resources.

The goal of this thesis is therefore to find methods of converting successful prompts in a resource-rich source language A to perform classification tasks (sentiment/emotion analysis, etc) in an under-resourced language B.

Prerequisites: knowledge of methods in deep learning, experience in a natural language understanding task, for instance emotion classification, named entity recognition, or something else.

BSc/MSc: MSc

Keywords: Cross-lingual transfer, language models, classification, prompt based learning