

Announcement

02.05.2019

BACHELOR THESIS:

Subjective Evaluation of Text Readability, Understandability and Complexity

DESCRIPTION

Text is a major medium for transforming information in daily human to human communication. *Text Readability* is described as the sum of all text elements that affect the readers understanding, reading speed and level of interest in the material. A close concept is *Text Complexity* which is defined as a metric that determines how challenging a text is for a reader. Due to the major role of text in day to day communication, *text simplification* is becoming a popular task in Natural Language Processing. The text simplification is a process of translating a given text from one language to the same language but making it simpler to understand given a specific target group. Texts with uncommon words, long sentences and complicated grammar structure can be difficult to read and understand for many people, and machines. Especially for people with poor literacy, cognitive/linguistic impairment, or those with language learners. A key question in text simplification research is to estimate the text readability, complexity and understandability. These measures can be used in the *evaluation* process in which quality of different (automatic/manual) text simplification processes can be estimated. To trained prediction models, annotated dataset is needed.

In this thesis, the candidate should conduct subjective tests to further develop an existing dataset with text readability, understandability and complexity score for German language. The test participants must belong to language learner groups in level B1. The candidate needs to study the state-of-the-art approach on how to subjectively measure those concepts: Furthermore he/she should design and conduct the subjective tests in the laboratory.

Text analysed in this thesis will be in German, however the thesis should be written in English.

REQUIREMENTS

- HTML/CSS/JS programming skills
- Experience with conducting subjective tests (Usability Engineering)
- Speaking German
- Linguistics knowledge is a plus
- Great interest to the topic and well organized

CONTACT

Dr. Babak Naderi, babak.naderi@tu-berlin.de
(Deadline: 31.05.2019)