

Supervisor assessment of Bachelor's Thesis

Student: Vašš Adam
Title: Topic Identification from Spoken TED-Talks (id 21519)
Supervisor: Kesiraju Santosh, UPGM FIT VUT

1. Assignment comments

The work is considered more than a moderately difficult as it involves understanding concepts in building ASR systems and training topic identification classifiers. Although there are open source tools available for these tasks, it requires some amount of understanding on how to use them appropriately. The task also involves creating topic ID annotations for TED talks by aggregating the tags.

2. Literature usage

The work involves study of (1) basic building blocks of ASR systems and (2) topic ID classifiers. The literature study of ASR related concepts is rather unsatisfactory whereas studies regarding topic ID classifiers was satisfactory.

3. Assignment activity, consultation, communication

The student was moderately active during the beginning of the project, and was more active towards the end. The student was well prepared for the consultations. The student missed few deadlines in the first semester but managed to do well in the second semester.

4. Assignment finalisation

The final work was partly incomplete as the student came across few problems dealing with the Kaldi toolkit for building ASR systems. There was not enough time to resolve them as most of the work was done during the end of the semester. On the other hand, the student has done considerable work in preparing annotations for TED talks and training topic ID classifiers.

5. Publications, awards

One of the possibilities of the work is to release (in the form of publication) topic annotations and baseline systems for TED talks. However, since the work was not 100% complete, it is not possible to do so.

6. Total assessment

satisfactory (D)

Satisfied with:

1. Analysis of data and preparation of topic ID annotations for TED talks.
2. Understanding of topic ID classifiers.
3. Implementation of baseline topic ID systems.

Partly unsatisfied with:

1. Literature survey regarding ASR systems.
2. Fully decoding the test recordings (TED talks) using the trained ASR system.
3. Work done was more than what was written in the thesis.

In Brno 30. May 2019

.....
signature