



Supervisor's statement of a final thesis

Student: BA. Pragalbha Lakshmanan
Supervisor: Ing. Milan Dojčinovski, Ph.D.
Thesis title: Enrichment of the DBpedia NIF dataset
Branch of the study: Software Engineering

Date: 8. 6. 2020

<i>Evaluation criterion:</i>	<i>The evaluation scale: 1 to 4.</i>
1. Fulfilment of the assignment	<u>1 = assignment fulfilled,</u> 2 = assignment fulfilled with minor objections, 3 = assignment fulfilled with major objections, 4 = assignment not fulfilled
<i>Criteria description:</i> Assess whether the submitted FT defines the objectives sufficiently and in line with the assignment; whether the objectives are formulated correctly and fulfilled sufficiently. In the comment, specify the points of the assignment that have not been met, assess the severity, impact, and, if appropriate, also the cause of the deficiencies. If the assignment differs substantially from the standards for the FT or if the student has developed the FT beyond the assignment, describe the way it got reflected on the quality of the assignment's fulfilment and the way it affected your final evaluation.	
<i>Comments:</i> The main goal of the thesis was to enable text pre-processing of the DBpedia NIF dataset. The student implemented a CLI tool which can be used to execute selected NLP techniques over the DBpedia NIF dataset in different languages. The student validated and evaluated the results on a subset of data. The main objectives of the thesis have been fulfilled.	
<i>Evaluation criterion:</i>	<i>The evaluation scale: 0 to 100 points (grade A to F).</i>
2. Main written part	68 (D)
<i>Criteria description:</i> Evaluate whether the extent of the FT is adequate to its content and scope: are all the parts of the FT contentful and necessary? Next, consider whether the submitted FT is actually correct – are there factual errors or inaccuracies? Evaluate the logical structure of the FT, the thematic flow between chapters and whether the text is comprehensible to the reader. Assess whether the formal notations in the FT are used correctly. Assess the typographic and language aspects of the FT, follow the Dean's Directive No. 26/2017, Art. 3. Evaluate whether the relevant sources are properly used, quoted and cited. Verify that all quotes are properly distinguished from the results achieved in the FT, thus, that the citation ethics has not been violated and that the citations are complete and in accordance with citation practices and standards. Finally, evaluate whether the software and other copyrighted works have been used in accordance with their license terms.	
<i>Comments:</i> The thesis is well organised and structured in chapters and sections which logically follow. All necessary parts are present and described. The thesis is well formatted (except for the issues indicated below). The citations ethics have not been violated and the citations are complete in accordance with the citation norms. There are, however, several aspects of the thesis that could be improved. <ul style="list-style-type: none">- The related work Section 2.2. discusses minimal number of related works and the thesis is not well positioned wrt the related work. The related work is not well summarised and the gaps clearly identified.- The results from the experiments are not summarised and the main findings are not explicitly indicated.- The conclusion section does not clearly summarise the thesis and the main results.- The formatting of the listings could be improved.- The language and grammar is satisfactory, however, some parts are difficult to read and follow.	
<i>Evaluation criterion:</i>	<i>The evaluation scale: 0 to 100 points (grade A to F).</i>
3. Non-written part, attachments	90 (A)
<i>Criteria description:</i> Depending on the nature of the FT, comment on the non-written part of the thesis. For example: SW work – the overall quality of the program. Is the technology used (from the development to deployment) suitable and adequate? HW – functional sample. Evaluate the technology and tools used. Research and experimental work – repeatability of the experiment.	

Comments:

The student implemented a CLI tool in Python for execution of several NLP tasks.
The tool supports several languages and integrates several NLP tools.
The tool can be configured for the language, data and NLP tool.
The technology used is suitable and adequate.
The tool is documented and the codebase published on github.

While most parts are automated, some parts still require manual effort, e.g. download of data. This step could be also automated.

Evaluation criterion:

The evaluation scale: 0 to 100 points (grade A to F).

4. Evaluation of results, publication outputs and awards

80 (B)

Criteria description:

Depending on the nature of the thesis, estimate whether the thesis results could be deployed in practice; alternatively, evaluate whether the results of the FT extend the already published/known results or whether they bring in completely new findings.

Comments:

The tool is functional, however, its deployment in practice, in real-world environment, is limited and would require more effort.

The work and the results from the thesis are still not ready for a publication since some parts require improvement. Primarily, the experiments could be more detailed and better performed.

Evaluation criterion:

The evaluation scale: 1 to 5.

5. Activity and self-reliance of the student

5a:
1 = excellent activity,
2 = very good activity,
3 = average activity,
4 = weaker, but still sufficient activity,
5 = insufficient activity
5b:
1 = excellent self-reliance,
2 = very good self-reliance,
3 = average self-reliance,
4 = weaker, but still sufficient self-reliance,
5 = insufficient self-reliance.

Criteria description:

From your experience with the course of the work on the thesis and its outcome, review the student's activity while working on the thesis, his/her punctuality when meeting the deadlines and whether he/she consulted you as he/she went along and also, whether he/she was well prepared for these consultations (5a). Assess the student's ability to develop independent creative work (5b).

Comments:

The student consulted the work with the supervisor and the feedback received was considered in every subsequent iteration.

The student was always prepared for the consultation.

The thesis has been actively consulted especially in the final phase before the thesis submission.

The student has shown skills to develop an independent creative work, i.e. software, and get familiar with new, previously unknown technologies.

Evaluation criterion:

The evaluation scale: 0 to 100 points (grade A to F).

6. The overall evaluation

75 (C)

Criteria description:

Summarize which of the aspects of the FT affected your grading process the most. The overall grade does not need to be an arithmetic mean (or other value) calculated from the evaluation in the previous criteria. Generally, a well-fulfilled assignment is assessed by grade A.

Comments:

The assignment of the thesis has been fulfilled.

The student implemented a functional software which can be used to pre-process and enrich the DBpedia NIF dataset with additional linguistic information.

There are several issues with the main written part of the thesis (motivation, related work, experiments and conclusion), which however, do not have a significant impact on the final quality of the thesis.

The student has managed to apply the knowledge acquired during the studies and implement a functional software.

Considering the comments above, I recommend grade C.

Signature of the supervisor: