



Supervisor's statement of a final thesis

Student: Vladimir Ananyev
Supervisor: Ing. Tomáš Borovička
Thesis title: Approximation Algorithms for Clustering
Branch of the study: Computer Science

Date: 11. 2. 2019

<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 thesis fulfilled the assignment completely without any objection.	
<i>Evaluation criterion:</i>	<i>The evaluation scale: 0 to 100 points (grade A to F).</i>
2. Main written part	90 (A)
<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 well explains the problem of clustering and particularly agglomerative hierarchical clustering. Student addresses the problem of computational complexity and prepares motivation for approximative methods that follow. The thesis is logically structured, and it is easy to follow. The text is to the point and there are not any unnecessary parts. The language is on a good level. Student cites relevant literature according to the citation standards.	
<i>Evaluation criterion:</i>	<i>The evaluation scale: 0 to 100 points (grade A to F).</i>
3. Non-written part, attachments	80 (B)
<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.	
<i>Comments:</i> Student implemented LSH-link and MST-linkage algorithms in c++. The quality of the code and documentation are without any objection. However, the experimental part could be more extensive and the results should be more thoroughly explained.	
<i>Evaluation criterion:</i>	<i>The evaluation scale: 0 to 100 points (grade A to F).</i>
4. Evaluation of results, publication outputs and awards	90 (A)
<i>Criteria description:</i> 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 thesis provides an introduction to approximate methods of agglomerative hierarchical clustering. The experiment demonstrates the scalability of the LSH-link method and compares it with the exact method on several datasets. The results show that the approximate methods can provide similar results with significantly lower time complexity. The work is based on facts published in the literature. The implementation of LSH-link is a valuable outcome as the implementation was not available at the time of writing the thesis.

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 was very active, always prepared for consultation and in general, he was dealing with the thesis in a responsible manner. I was satisfied with the collaboration and delivered results.

Evaluation criterion:

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

6. The overall evaluation

85 (B)

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 thesis is well written. The student was communicative and responsible when writing the thesis. The theoretical part is comprehensive, simple to read and follow. The experimental part is, however, a bit weaker. I recommend the thesis to be classified with grade B.

Signature of the supervisor: