FACULTY OF INFORMATION TECHNOLOGY CTU IN PRAGUE

Supervisor's statement of a final thesis

Student:	Matúš Žilinec
Supervisor:	doc. Ing. Pavel Kordík, Ph.D.
Thesis title:	Question Answering Algorithms in Natural Language
Branch of the study:	Knowledge Engineering

Date: 16. 6. 2019

Evaluation criterion:	The evaluation scale: 1 to 4.
1. Fulfilment of the assignment	***1 = assignment fulfilled,*** *2 = assignment fulfilled with minor objections,* *3 = assignment fulfilled with major objections,* *4 = assignment not fulfilled*

Criteria description:
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.

Comments:
The thesis fulfills the assignment.

Evaluation criterion:	The evaluation scale: 0 to 100 points (grade A to F).
2. Main written part	*100 (A)*

Criteria description:
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.

Comments:
Text is very informative and well written. I enjoyed reading the theoretical part - it is apparent that author is well oriented in the field and knowledgeable of modern deep learning approaches such as the Transformer or BERT. The experimental part is also excellent. I really like the analysis of data and errors on difficult QA pairs.

Evaluation criterion:	The evaluation scale: 0 to 100 points (grade A to F).
3. Non-written part, attachments	*99 (A)*

Criteria description:
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:
Implemented system is first BERT based QA network for the NaturalQuestions dataset capable of running on standard GPUs.

Evaluation criterion:	The evaluation scale: 0 to 100 points (grade A to F).
4. Evaluation of results, publication outputs and awards	*100 (A)*

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:
We plan to extend the results into a journal publication.

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:

Matus works independently, he is always prepared.

Evaluation criterion:	*The evaluation scale: 0 to 100 points (grade A to F).*
6. The overall evaluation	*99 (A)*

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:

Excellent thesis on the edge of state of the art in the important field of natural language understanding.

Signature of the supervisor: