



Supervisor's statement of a final thesis

Student: Alex Eduard Marksfeld
Supervisor: Ing. Mgr. Jan Romportl, Ph.D.
Thesis title: Strojové učení v sociodemografické segmentaci zákazníků telekomunikační společnosti
Branch of the study: Knowledge Engineering

Date: 11. 6. 2018

<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> All the objectives of the FT were successfully reached. The student has developed the FT beyond the assignment by implementing and testing a convolutional neural network on gender prediction using Call Detail Record data.	
<i>Evaluation criterion:</i>	<i>The evaluation scale: 0 to 100 points (grade A to F).</i>
2. Main written part	75 (C)
<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 logical structure of the FT is adequate, providing a brief general introduction to supervised ML methods, more specific introduction to the tools and data available in the Telco company, and detailed information about how the student proposed and tested his predictive models. The information flow across the chapters is usually good, but there are some surprising wobbles present from time to time. Unfortunately, the text quality is very disbalanced across the FT. There are sections where the reader must struggle quite hard to understand what the author wanted to deliver – and still I think somewhere the information is delivered somewhat distorted to a reader who does not have a proper background and the level of detail about the Telco's data. The same holds for English too: some paragraph are written very well, whereas there are quite a few hardly comprehensible places. Other formal aspects of the FT (typesetting, figures, citations, biblio, etc.) are OK.	
<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.	
<i>Comments:</i> The scripts both for data preparation/wrangling and for predictive modelling are of a very good quality. They are directly reusable in other projects of the student's peers working with the same data.	
<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	80 (B)
<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 results in a form of the scripts and trained models can definitely be deployed in practice. The FT extends already published results, but on untested data and integrating quite a few other sources. Moreover, the paper that brought the inspiration to test CNNs on CDR, is only from November 2017, which means that the student worked within the state of the art. So I would like give A in this category but unfortunately I cannot because of the writing part of the FT: as mentioned above, it often fails to deliver the results properly to a reader, deteriorating the quality of the results of the FT.

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 well prepared for all the meetings and discussions. He was very smart, proactive and self-reliant in working with the data, algorithms and also the other members of the data science team. From the point of a thesis supervisor, it was very pleasant to cooperate with such a student. However, I was eventually quite surprised how much help he needed in creating a coherent and well readable piece of text (and still it did not end up so well). Sometimes I am wondering whether this is not a doom of the whole current cohort of the students: they are excellent in doing the analytical job but when it comes to writing it up, they tend to fail.

Evaluation criterion:

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

6. The overall evaluation

77 (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:

Here I am going to take an opportunity to be tough on Alex: I had very high expectations of his work; and most importantly, I believe he himself had expected he would deliver the written thesis in much higher quality. He started very well but then he did not have enough time or energy to finish properly. Due to the fact how well he worked with the data and how dutifully and proactively he cooperated in the last year, I am tempted to give him B. However, considering how intelligent, hard-working and also goal-directed he is, I think it is imperative for him to do much better next time. Therefore, as a token of reprehension and reminder to work more thoroughly to match the expectations set up by his intelligence and world views, I am rather presenting him with overall C.

Signature of the supervisor: