Supervisor's statement of a final thesis

Czech Technical University in Prague
Faculty of Information Technology

Student: Bc. Jan Veselý
Supervisor: Ing. Jan Černý
Thesis title: Analýza chování modelů strojového učení
Branch of the study: Knowledge Engineering

Date: 2. 6. 2017

1. Difficulty and other comments on the assignment

   Criteria description:
   Characterize this final thesis in detail and its relationships to previous or current projects. Comment what is difficult about this thesis (in case of a more difficult thesis, you may overlook some shortcomings that you would not in case of an easy assignment, and on the contrary, with an easy assignment those shortcomings should be evaluated more strictly.)

   Comments:
The difficulty of this assignment lies in the ability to be able to design and implement things in a general way so that it can work for any type of models and also for multiple use cases from automated model testing between different code versions to performance evaluation. Difficulty from the algorithmic point of view is average.

2. Fulfilment of the assignment

   Criteria description:
   Assess whether the thesis meets the assignment statement. In Comments indicate parts of the assignment that have not been fulfilled, completely or partially, or extensions of the thesis beyond the original assignment. If the assignment was not completely fulfilled, try to assess the importance, impact, and possibly also the reason of the insufficiencies.

   Comments:
   All goals of the assignment were addressed and completed in the thesis.

3. Size of the main written part

   Criteria description:
   Evaluate the adequacy of the extent of the final thesis, considering its content and the size of the written part, i.e. that all parts of the thesis are rich on information and the text does not contain unnecessary parts.

   Comments:
   I would like to see more research around metrics for model performance measuring and more detailed analysis of this area, all other parts of the thesis are described adequately.

4. Factual and logical level of the thesis

   Criteria description:
   Assess whether the thesis is correct as to the facts or if there are factual errors and inaccuracies. Evaluate further the logical structure of the thesis, links among the chapters, and the comprehensibility of the text for a reader.

   Comments:
   Thesis is comprehensible and easily understandable.

5. Formal level of the thesis

   Criteria description:
   Assess the correctness of formalisms used in the thesis, the typographical and linguistic aspects, see Dean’s Directive No. 14/2015, Article 3.

   Comments:
   I don’t have any major objections on the form of the thesis.

6. Bibliography

   Criteria description:
   Evaluate the student’s activity in acquisition and use of studying materials in his thesis. Characterize the choice of the sources. Discuss whether the student used all relevant sources, or whether he tried to solve problems that were already solved. Verify that all elements taken from other sources are properly differentiated from his own results and contributions. Comment if there was a possible violation of the citation ethics and if the bibliographical references are complete and in compliance with citation standards.
**Comments:**
There are only few sources mentioned in the thesis and research part of the metrics could be more detailed.

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

<table>
<thead>
<tr>
<th>7. Evaluation of results, publication outputs and awards</th>
<th>85 (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criteria description:</td>
<td>Comment on the achieved level of major results of the thesis and indicate whether the main results of the thesis extend published state-of-the-art results and/or bring completely new findings. Assess the quality and functionality of hardware or software solutions. Alternatively, evaluate whether the software or source code that was not created by the student himself was used in accordance with the license terms and copyright. Comment on possible publication output or awards related to the thesis.</td>
</tr>
</tbody>
</table>

**Comments:**
Result of the thesis is a well designed and well programmed system for analysing behaviour of machine learning models. The system is designed as a general framework which allows to use it for different type of models such as prediction models which works with model matrix or behavioral models working with transactional data. System is also easily scalable and adding new metrics is very easy due to well defined interfaces and project structure.

Second goal was also met by implementing baseline metrics that are most commonly used and also implementing metrics to monitor changes in models prediction with new model deployment. This is used as advanced testing layer to discover bugs before releasing new model source code to the production.

Only downside is that those baseline metrics does not perform well enough on some use cases and there is certainly a space for future work regarding improving those or implementing new ones.

**Evaluation criterion:**
No evaluation scale.

<table>
<thead>
<tr>
<th>8. Applicability of the results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criteria description:</td>
</tr>
</tbody>
</table>

**Comments:**
Whole system currently runs in production and is processing thousands of jobs per day and sending tens of millions of requests per day to the running models.

**Evaluation criterion:**
The evaluation scale: 1 to 5.

<table>
<thead>
<tr>
<th>9. Activity and self-reliance of the student</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criteria description:</td>
</tr>
<tr>
<td>Comments:</td>
</tr>
<tr>
<td>9a:</td>
</tr>
<tr>
<td>9b:</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>10. The overall evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criteria description:</td>
</tr>
<tr>
<td>Comments:</td>
</tr>
<tr>
<td>10</td>
</tr>
</tbody>
</table>

Signature of the supervisor: