



Supervisor's statement of a final thesis

Supervisor: Ing. Adam Valenta
Student: Bc. Yuliia Syzon
Thesis title: Standard Decision Trees in Machine Learning
Branch / specialization: Knowledge Engineering
Created on: 5 February 2023

Evaluation criteria

1. Fulfillment of the assignment

- ▶ [1] assignment fulfilled
- [2] assignment fulfilled with minor objections
- [3] assignment fulfilled with major objections
- [4] assignment not fulfilled

The assignment is fulfilled in its full scope. The scope of the assignment is adequate for the diploma thesis, and the quality of the work is very good. The student correctly defines objectives, explain all the terminology used in the text (section Preliminaries), breakdown Decision Trees and try to find common parts among the various building methods (design steps for Decision Tree), and use Software Engineering process for the implementation of SDT (section Requirements and Evaluation). The student also explains some parts of the theory on the example data, which is handy for an Engineer in implementation.

2. Main written part

90 /100 (A)

I appreciate that the thesis is written in English, uses the new faculty template, and is written in the formal but legible language. The student correctly put the SDT into the general Machine Learning field, explored boundaries with similar means, and provided Mathematical and engineering background for provided methods. Most of the parts of the thesis logically and thematically follow up, but some appeared from nowhere and would be good to have as an extension. Overall the student used relevant sources that were quoted correctly with only minor objections. The implementation part could also review a current decision tree implementation and define its pros and cons. A review of the current implementation would connect better theory and implementation itself. Evaluation would also benefit from wider datasets for prediction performance, to see if the performance is still adequate, and for scalability performance, where it would be good to ensure that with more CPU, the time correctly descent. I also recommend analyzing how quickly prediction converges compared to other implementations. I

appreciate that student overall creates her own Figures that precisely fit the written topic.

3. Non-written part, attachments

95 /100 (A)

The implementation part follows H2O.ai's standards, and the student correctly defines what needs to be improved and provides unit tests to prove that implementation is correct. What needs to be improved is the internal API of SDT, which does not follow the open-close principle.

4. Evaluation of results, publication outputs and awards

98 /100 (A)

The current implementation and quality standards H2O.ai's review process are met. The student is going to cooperate with H2O.ai open-source team and AutoML to finish Drawbacks and future improvements to merge the production version of SDT.

5. Activity of the student

- ▶ [1] **excellent activity**
- [2] very good activity
- [3] average activity
- [4] weaker, but still sufficient activity
- [5] insufficient activity

Student proactively consulted all faced issues, was motivated, and finished all the assigned work by the deadline.

6. Self-reliance of the student

- ▶ [1] **excellent self-reliance**
- [2] very good self-reliance
- [3] average self-reliance
- [4] weaker, but still sufficient self-reliance
- [5] insufficient self-reliance

The student is as reliable as every other member of H2O.ai's team.

The overall evaluation

92 /100 (A)

The student demonstrated the ability to understand a technical-scientific problem and implement this problem into an actual application. The result is a contribution to the open-source library, a huge codebase with much integration, which is hard to start to contribute. I appreciate the activity and reliability of the student's overall thesis quality with a challenging assignment.

Instructions

Fulfillment of the assignment

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.

Main written part

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. 52/2021, 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.

Non-written part, attachments

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.

Evaluation of results, publication outputs and awards

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.

Activity of the student

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.

Self-reliance of the student

From your experience with the course of the work on the thesis and its outcome, assess the student's ability to develop independent creative work.

The overall evaluation

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.