



Supervisor's statement of a final thesis

Supervisor: Ing. Dominik Soukup
Student: Bc. Lukáš Jančíčka
Thesis title: Concept drift and model degradation in network traffic classification
Branch / specialization: Knowledge Engineering
Created on: 18 January 2024

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 master thesis is focused on concept drift analysis in the computer network domain. The result is a detailed analysis of network traffic behavior and existing methods for concept drift detection. The results also include a software prototype that identifies shifts in the underlying distribution of network traffic. The quality of this work is above normal standards and contains parts above the assignment (best practices for ML deployment in computer network traffic environment, possibility to run the prototype for offline or online dataset)

2. Main written part 100/100 (A)

The written part is very well structured and easy to follow. Language is also correctly used even though English is not the student's native language. Typography and citations are accurate, rich enough, and consistent throughout the work. It's notable that the student continuously worked on this thesis and did not complete it at the last minute.

3. Non-written part, attachments 100/100 (A)

The non-written part contains the required software prototype and also several Jupyter notebooks. Overall, the non-written part is at a very good level and ready even for deployment in more production environments.

4. Evaluation of results, publication outputs and awards

100/100 (A)

The results achieved are very valuable for network traffic analysis and deployment of machine learning models. The knowledge is based on the current state of the art, therefore, the implemented prototype could be beneficial in practice. This work also includes a novel approach of weighing the drift result by the feature importances, therefore, there is very promising potential in future conference publication.

5. Activity of the student

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

The student's activity was at an excellent level. He actively asked questions and regularly visited check-in meetings where he was always prepared.

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 started exploring this topic in 2022 and was very curious during he whole time. The development of the software prototype was not easy since it was related to many existing tools and frameworks. Regardless of the challenges, the student delivered results above the initial assignment.

The overall evaluation

100/100 (A)

This master thesis completely achieves the required assignment and delivers even the findings above. The student had to learn many new frameworks and study the existing state of the art. The outcome is a concept drift detector and behavior analysis of QUIC and TLS protocols that have not been done in related work. The implemented detector was validated on real datasets, and it is ready for production. Moreover, there is good potential for conference paper.

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.