

I. IDENTIFICATION DATA

Thesis title:	Comparing Methods for Detecting ConceptDrift in Data Streams
Author's name:	Daniil Barabašev
Type of thesis :	bachelor
Faculty/Institute:	Faculty of Electrical Engineering (FEE)
Department:	Department of Cybernetics
Thesis reviewer:	doc. Mgr. Viliam Lisý, MSc., Ph.D.
Reviewer's department:	Department of Computer Science

II. EVALUATION OF INDIVIDUAL CRITERIA

Assignment	ordinarily challenging
<i>How demanding was the assigned project?</i>	
The student had to find, implement, and compare techniques from research literature. These techniques and the problem they solve were not covered in the student's studies. On the other hand, the problem and techniques are not conceptually more complex than the common content of bachelor studies.	

Fulfilment of assignment	fulfilled
<i>How well does the thesis fulfil the assigned task? Have the primary goals been achieved? Which assigned tasks have been incompletely covered, and which parts of the thesis are overextended? Justify your answer.</i>	
All points of the assignment are fulfilled reasonably well, but none of them is substantial overachieved. The student did very solid necessary work, but not much extra.	

Activity and independence when creating final thesis	C - good.
<i>Assess whether the student had a positive approach, whether the time limits were met, whether the conception was regularly consulted and whether the student was well prepared for the consultations. Assess the student's ability to work independently.</i>	
The student was exceptionally independent. He was coming up with his own materials and approaches, and he did vast majority of writing completely independently as well. He was generally prepared and on time for the consultations. However, the consultations rarely took more than 15 minutes and the student never had much to ask and discuss. There was some lack of communication in the very beginning of the work and the final text of the thesis was not completed early enough for me to give a proper comprehensive feedback.	

Technical level	C - good.
<i>Is the thesis technically sound? How well did the student employ expertise in his/her field of study? Does the student explain clearly what he/she has done?</i>	
The clarity of the contribution could have been improved in general. There are a few larger technical issues. For example, the student reports large part of the results on just one run of the algorithms with one random seed. Ideally, there should be more seeds, or it should be discussed in much more details why it may not be too big issue in this particular case. Another example is Section 2.2., where student makes too strong conclusions, without proper discussion, supported by only one related work. Also, the motivation for the student's extension of the methods found in the literature and some more decisions in the thesis are not very well explained. Still, the overall approach the student has chosen and the execution of the thesis was good and he demonstrated his expertise in the field.	

Formal level and language level, scope of thesis	D - satisfactory.
<i>Are formalisms and notations used properly? Is the thesis organized in a logical way? Is the thesis sufficiently extensive? Is the thesis well-presented? Is the language clear and understandable? Is the English satisfactory?</i>	
The text was written at the last minute. The structure of the text is not very clear. Review of the literature, introduction of algorithms and experimental results are interleaved. High-level descriptions of the algorithms for concept drift detection are at four places. Some set of algorithms is introduced initially, but a different set is later evaluated. Algorithms are	

described at one place, but their pseudocode comes only with the experiments. One new algorithm is introduced only at the experiments along with its explanation and pseudocode. Another algorithm introduced at the beginning (RBM-DD) is not mentioned anymore without clear explanation. The thesis would be much easier to read if it had a clearer structure. There are also some typos, use of terminology that has not been properly defined before, and unintuitive flow among consequent sections, e.g., at the beginning of Section 2.3.

Selection of sources, citation correctness

D - satisfactory.

Does the thesis make adequate reference to earlier work on the topic? Was the selection of sources adequate? Is the student's original work clearly distinguished from earlier work in the field? Do the bibliographic citations meet the standards?

This is, in my opinion, the weakest point of the thesis. On page 8, the thesis refers to reference [10] for details of an OLINDDA algorithm, but I did not find such algorithm in the paper. Figure 2.4. is redrawn almost one-to-one from prior work, but it is not clearly acknowledged. The bibliography section is inconsistent and the records are incomplete. For example, [5,7,15] are missing the publication, where the paper was published. [9] and some others have proper journal version, but only arXiv is referenced. [16] includes editors, full date, URL and other data which is missing for majority of other references.

Additional commentary and evaluation (optional)

Comment on the overall quality of the thesis, its novelty and its impact on the field, its strengths and weaknesses, the utility of the solution that is presented, the theoretical/formal level, the student's skillfulness, etc.

Please insert your comments here.

III. OVERALL EVALUATION, QUESTIONS FOR THE PRESENTATION AND DEFENSE OF THE THESIS, SUGGESTED GRADE

Summarize your opinion on the thesis and explain your final grading.

Question:

- Why is it sufficient to use only one random seed to the experiments presented in the thesis?

The student worked very independently and consistently, and fulfilled the assignment of the thesis in all points. There were some issues with structure of the text, clarity of explanations of some decisions, and references, but overall the thesis demonstrates the student's ability to understand a new technical problem, find relevant literature and implement the proposed solutions.

The grade that I award for the thesis is **C - good**.

Date: **4.6.2024**

Signature: