THESIS REVIEWER'S REPORT

I. IDENTIFICATION DATA

Thesis title:	**Reservoir Computing Framework v Apache Flink**
Author's name:	**Hynek Noll**
Type of thesis :	bachelor
Faculty/Institute:	Faculty of Electrical Engineering (FEE)
Department:	Department of Computer Science
Thesis reviewer:	Ing. Jan Drchal, PhD.
Reviewer's department:	Department of Computer Science

II. EVALUATION OF INDIVIDUAL CRITERIA

Assignment **challenging**

How demanding was the assigned project?

The complexity of the assignment depends on the level of parallelism implemented. It was not directly stated in the assignment; however, the technologies involved need significant time to master.

Fulfilment of assignment **fulfilled**

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.

The assignment was fulfilled with no objections.

Methodology **partially applicable**

Comment on the correctness of the approach and/or the solution methods.

It is not clear to me why so large part of the work focuses on the implementation of linear regression: a quick internet search reveals that it is already part of the Apache Flink codebase. The implemented methods are tested on very few and, more importantly, small datasets which is not in agreement with the intended use as Apache Flink is meant for processing big data. I have an important objection to the experimental part: 1) it is hardly reproducible as many details are missing, 2) datasets are insufficiently described (I miss the number of samples, split information, input dimension, etc.), 3) the evaluation of experiments is too brief, 4) the figures are not well described, and they are not referred from the text, 5) the linear regression part of the system should have been compared to other existing implementations.

Technical level **D - satisfactory.**

Is the thesis technically sound? How well did the student employ expertise in the field of his/her field of study? Does the student explain clearly what he/she has done?

The source code, as well as the text show, that student is skilled in software design and programming; on the other hand, the experimental part is very confusing as mentioned above.

Formal and language level, scope of thesis **C - good.**

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?

The text of the work is written exceptionally well, and up to section 5.5 - Implementation Description it was really a pleasure to read. While the text of 5.5 is still comprehensive, I miss important details on system architecture. It is not clear at which level and how the parallelization is realized by Apache Flink. What modules (operators) are used and what kind of data flows between them. A block diagram would be revealing. The worst part of the text is definitely the section describing the experiments as described above.

Selection of sources, citation correctness **A - excellent.**

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?

No objections.

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.

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. Pose questions that should be answered during the presentation and defense of the student's work.

While the textual part of the work is written exceptionally well, I miss a better description of the software architecture. Final experiments lack comparison to other implementations, more realistic datasets, and, most importantly, performance and scalability assessment.

The grade that I award for the thesis is **D - satisfactory.**

Date: **3.6.2020** Signature: