

I. IDENTIFICATION DATA

Thesis title:	Survey of ML Model Serving Solutions and criteria for selection thereof
Author's name:	Petr Poliak
Type of thesis :	Master Thesis
Faculty/Institute:	Faculty of Electrical Engineering
Department:	Department of Computer science
Thesis reviewer:	Vojtěch Tůma
Reviewer's department:	Externist

II. EVALUATION OF INDIVIDUAL CRITERIA

Assignment	A / Excellent
<i>How demanding was the assigned project?</i>	
<p>The Author had to familiarise with a complex and rapidly changing topic, with no standard best-practices existing. This required autonomy and time investment, and also keen judgement -- quite a few existing information sources in this area are downright misleading.</p> <p>The objective / the research question of the Thesis was defined fairly loosely, with expectation having been put on the Author to show confidence and fill in the subject autonomously. The structure of the Thesis and presentation of the results have changed a few times during the work, testifying that the Author has been actively experimenting and optimising.</p> <p>The work on the Thesis required multiple implementation side projects to gain further understanding and evaluate some of the proposed concepts -- and as those are not part of the Thesis, it is not obvious to the uninitiated reader.</p>	

Fulfilment of assignment	B / Very Good
<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>	
<p>The original assignment, as stated in the Thesis definition, has been pretty much fulfilled. The Reviewer is only pointing out firstly, that the Evaluation section is containing only two, not three cases -- the third case study has actually been carried out and written by the Author, but was not allowed to be published. Secondly, the Evaluations are rather retro-fitting descriptive analyses using Author's framework, they don't contain any pro-active findings or inferential outcomes, to be compared with decisions made by Experts.</p>	

Methodology	A / Excellent
<i>Comment on the correctness of the approach and/or the solution methods.</i>	
<p>The Reviewer positively notes how the Author decomposed the problem in a methodologically sound way -- first defining Evaluation Criteria for the Thesis Objective, then postulating the Objective in a way of Model Deployment Approaches, and lastly, giving Application Criteria to ensure smooth applicability of the Thesis.</p> <p>Furthermore, the Author correctly included a wider Survey (with smallness of the surveyed population resulting from the specificity of the problem, not from the Author's lack of field work), and subjected his results to Expert critique often during the work.</p> <p>Lastly, the Author validated some of the results by actual implementations, not relying on presumptions.</p>	

Technical level	A / Excellent
<i>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?</i>	
<p>The Reviewer is aware of all the side projects to support the Thesis' outcome that have been done during the Author's work on the Thesis, and those exhibited a great level of technical excellence. A certain flaw may be that the Thesis itself does not refer to those, but this was an early and sober decision during scoping of the Thesis.</p> <p>All in all, the Reviewer has not found any technical weakness in the Thesis.</p>	

Formal and language level, scope of thesis	B / Very 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 presentation and the language are well done, the Thesis is a very accessible introduction to the topic as a whole even to the non-Expert. The Reviewer has not identified any formal or notational flaw.

Only presentational downside may be that the Introduction is not sufficiently motivated -- a naive reader may be tempted to believe the promise of certain existing solutions being marketed by big players, without realising all the downsides and consequences. This is to an extent addressed later in the text -- but some sort of story to motivate the Thesis as whole would help the mission. Not including a proper motivation for the non-Expert is a usual bias of those well-versed in the subject.

Another downside, possibly in the area of extensibility or logicallity, is that the chapter which gives a methodology on how to apply the main result, could perhaps be structured differently -- instead of saying what is in effect "if your use case is small, use a simple solution; if it is large, use a complicated one", the presentation could have gone via "when it is time to leave the simple solution for a complex one, and what are the best practices to follow when implementing the complex one.", or something like "transition complexity matrix". This would have been compatible and coherent with the previous chapter, and perhaps provide a more practical or logical outcome.

Lastly, some sort of compact comparison of the Model Deployment Schemata is missing -- something in spirit of page 11 (beginning of the Deployment Scheme Definition chapter), but in the context of the Deployment Scheme Selection chapter -- this would also improve the Thesis applicability and presentation.

Level of English is sufficient and gives a nice read.

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?

The Author has gone out and selected a sufficient number of earlier work. Of those, all are sound and objective. As noted here in the Assignment part, this has been particularly complicated, because of the newness of the Area, because of its lying in between multiple domains (ML, Operations, SW, ...), because of many sources that are rather marketing than scientific.

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.

The Reviewer finds possible additional value in the Evaluation Criteria chosen for the Thesis Objective -- those are perhaps of wider applicability than just in here (say, in a few years, when we'll be evaluating Quantum-Computing Model Deployment options).

None of the comments given here points out any serious flaw in the Thesis -- rather, they should be taken as suggestions for future work. This is a vast topic with many open questions, technical and methodological, and such are prone to overly ambitious Scopes where many rabbits are chased but none really caught. Here, the Author managed to avoid the trap and provided a compact self-sufficient outcome.

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.

The grade that I award for the thesis is **A / Excellent**



THESIS REVIEWER'S REPORT

Date: 30.05.2021

Signature: Vojtěch Tůma