

Jury Member Report – Doctor of Philosophy thesis.

Name of Candidate: Denis Kuznedelev

PhD Program: Computational and Data Science and Engineering

Title of Thesis: Accurate and efficient methods for neural network compression

Supervisor: Professor Dmitry Yarotsky

Name of the Reviewer: Max Ryabinin

I confirm the absence of any conflict of interest	
---	--

(Alternatively, Reviewer can formulate a possible conflict)	
---	--

	Date: 15-11-2024
--	-------------------------

The purpose of this report is to obtain an independent review from the members of PhD defense Jury before the thesis defense. The members of PhD defense Jury are asked to submit signed copy of the report at least 30 days prior the thesis defense. The Reviewers are asked to bring a copy of the completed report to the thesis defense and to discuss the contents of each report with each other before the thesis defense.

If the reviewers have any queries about the thesis which they wish to raise in advance, please contact the Chair of the Jury.

Reviewer's Report

The dissertation presents a comprehensive overview of the candidate's research on neural network compression. It presents the published works in the order that is arranged in related subgroups, and the candidate has made an additional effort to broadly categorize the methods studied in each of the papers with an illustration. Moreover, the thesis contains a detailed description of the background that will be helpful even on its own to the reader unfamiliar with the area. After the introductory sections, the dissertation contains a complete description of all methods developed by the candidate, as well as a summary of the empirical studies that were conducted. Overall, the text is well-written and clearly structured.

The topic of the dissertation fully reflects its content: in the manuscript, there is a summary of multiple papers published by the candidate, all belonging to the area described by the topic and exploring it from different angles (sparsity, quantization, knowledge distillation). The methods are also highly relevant to the studied research areas: the candidate explores established paradigms of model compression that

are a subject of active study of the scientific community, investigating those paradigms with novel approaches that are nonetheless grounded in prior work.

While I am not an expert on the topic of neural network compression, I conducted brief background research which confirmed that the methods developed by Denis and described in the thesis achieved state-of-the-art results in their respective areas at the time of their publication. The scientific results presented in the dissertation meet the highest quality bar, as evidenced by their acceptance at the highest-ranking international conferences in Machine Learning (ICLR, ICML, and NeurIPS). Overall, the proposed methods are very clearly explained (with theoretical justifications where applicable), and the experiments are conducted in relevant setups that are well-accepted in academic work. All of the results have direct applications in the resource-efficient deployment of neural networks, and some of the works even address the questions of hardware efficiency for the proposed methods.

In summary, I believe that the dissertation definitely qualifies for the award of a PhD degree. The only concerns I have are mostly about the presentation of the work. First, it might be beneficial to offer more illustrations or simple examples of how the proposed methods work to make them easier to understand for readers not familiar with the area of neural network compression. Second, there are a few minor typos and grammar mistakes in the manuscript that do not make it harder to understand and should be easy to address after brief proofreading.

Provisional Recommendation

I recommend that the candidate should defend the thesis by means of a formal thesis defense

I recommend that the candidate should defend the thesis by means of a formal thesis defense only after appropriate changes would be introduced in candidate's thesis according to the recommendations of the present report

The thesis is not acceptable and I recommend that the candidate be exempt from the formal thesis defense