THESIS SUPERVISOR'S REPORT

I. IDENTIFICATION DATA

Thesis title:	Generative models for high energy physics measurements
Author's name:	Lukáš Viceník
Type of thesis :	master
Faculty/Institute:	Faculty of Electrical Engineering (FEE)
Department:	Department of Cybernetics
Thesis reviewer:	Boris Flach
Reviewer's department:	Department of Cybernetics

II. EVALUATION OF INDIVIDUAL CRITERIA

Assignment **challenging**

How demanding was the assigned project?

The task required to familiarise with foundations of high energy physics and CERN Atlas detector responses, to "translate" the physical simulation of detector responses to a machine learning task and to apply advanced deep learning methods for generating augmented data. Furthermore, the task required to validate the suitability of the augmented data for training signal/background classifiers.

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 student has completed all tasks and achieved the thesis objectives.

Activity and independence when creating final thesis **A - excellent.**

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.

Lukáš Viceník has shown a very active and well organised approach. All parts of his work were regularly consulted both from the machine learning perspective and from the high energy physics perspective. He gave a talk on his results at the 2024 spring meeting of the German Physical Society.

Technical level **B - very good.**

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?

The student applied his expertise to solve a challenging machine learning task. The presentation of the corresponding concepts is technically correct. There are however minor drawbacks in the presentation of advanced concepts

Formal level 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 thesis is well structured and explains all relevant concepts and their implementation. It provides the necessary background information so that it is comprehensible both for physicists and computer scientists. In my opinion it is nevertheless too long. The English is satisfactory except for minor language errors.

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 references and citations are adequate. The students contribution is clearly distinguished from previous work.

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 thesis proposes, trains and validates a latent variable model for generating Atlas detector responses. The first chapters give the necessary information about the underlying physics and the CERN Atlas detector and introduce the classification concepts used for interpreting detector events. The next chapter introduces standard and hierarchical Variational autoencoders (VAE) and discusses two learning approaches for them. This concludes the theoretical part of the thesis. The second part of the thesis discusses the data preparation, the model architecture choices and the implementation of the proposed approach as well as the used evaluation metrics. The last part of the thesis presents the experiments and a through evaluation of the approach. This includes both an extensive comparison of generated and simulated events as well as an evaluation of the utility of generated data for training signal/background classifiers.

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

The student has clearly shown his ability to apply modern concepts of machine learning for a challenging application. This includes a clear understanding of the application, the adaptation of the machine learning concepts to it, as well as extensive experiments and a through validation of the achieved results.

The grade that I award for the thesis is **A - excellent.**

Date: **4.6.2024** Signature: