ČVUT
ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

SUPERVISOR'S OPINION OF FINAL THESIS

I. IDENTIFICATION DATA

Thesis name:	Axion-Like-Particle Search Using Machine Learning for the Signal Sensitivity Optimization with Run-2 LHC Data Recorded by the ATLAS Experiment
Author's name:	Ondrej Matoušek
Type of thesis :	bachelor
Faculty/Institute:	Faculty of Electrical Engineering (FEE)
Department:	Department of Cybernetics
Thesis supervisor:	Doc. Dr. André Sopczak
upervisor's department:	IEAP CTU in Prague

II. EVALUATION OF INDIVIDUAL CRITERIA

Assignment	**challenging**
For a bachelor thesis the topic has been very challenging, as besides the actual thesis tasks, the project required the student to get familiar with several software packages used at CERN in the ATLAS collaboration.	

Satisfaction of assignment	**fulfilled**
The thesis fulfills all assigned tasks.	

Activity and independence when creating final thesis	**A - excellent.**
The student showed great independence in fulfilling the tasks throughout the project.	

Technical Level	**B - very good.**
The student experience with the use of the provided data and applied his knowledge very well. The technical execution was straightforward. Testing of the software, and cross-checks of the results were performed. The documentation could have been extended to facilitate further use.	

Formal and language level, scope of thesis	**A - excellent.**
The thesis fulfills a high standard and is well structured.	

Selection of sources, citation correctness	**B - very good.**
Citations and resources are given in a correct and coherent format. Citation ethics has been fulfilled.	

Additional commentary and evaluation
The primary goal to apply machine learning for an analysis which used previously sequential cut selections was fulfilled. In the scope of a bachelor thesis best possible results were obtained. The technical realization was led to convincing results.

III. OVERALL EVALUATION, QUESTIONS FOR DEFENSE, CLASSIFICATION SUGGESTION

The efficiency of the selection was significantly improved over almost the whole mass range of the axion-like particle. The performance at the lowest mass was less than expected, and the student solved this problem by increasing the weight of the low mass in the training. He nicely demonstrated that his solution increased the performance also for the lowest mass. The thesis quality was high already in the first provided version. The student has been very attentive and performed consistently very well throughout the project. He presented his returns at group meetings regularly, and also presented preliminary results at the student session of the German Physical Society meeting in Dresden in March 2023.

I evaluate handed thesis with classification grade **A - excellent.**

Date: **7.6.2023** Signature:

REVIEWER'S OPINION OF FINAL THESIS

I. IDENTIFICATION DATA

Thesis name:	Axion-Like-Particle Search Using Machine Learning for the Signal Sensitivity Optimization with Run-2 LHC Data Recorded by the ATLAS Experiment
Author's name:	Ondřej Matoušek
Type of thesis :	bachelor
Faculty/Institute:	Faculty of Electrical Engineering
Department:	Department of Cybernetics
Thesis reviewer:	Boris Flach
Reviewer's department:	Department of Cybernetics

II. EVALUATION OF INDIVIDUAL CRITERIA

Assignment **ordinarily challenging**

Evaluation of thesis difficulty of assignment.

Overall, the level of difficulty of the task was normally challenging. On the one hand, the application domain required the student to become familiar with the basics of particle physics and the specifics of hadron collision detectors. On the other hand, the machine learning concepts required to solve the task, were moderately challenging.

Satisfaction of assignment **fulfilled**

Assess that handed thesis meets assignment. Present points of assignment that fell short or were extended. Try to assess importance, impact or cause of each shortcoming.

The submitted thesis meets the task assignments.

Method of conception **correct**

Assess that student has chosen correct approach or solution methods.

The proposed method is technically correct, however, with minor exceptions. It remains unclear for me, whether the used network architecture is already optimally chosen. The thorough analysis of the trained model is on the other hand very convincing.

Technical level **B - very good.**

Assess level of thesis specialty, use of knowledge gained by study and by expert literature, use of sources and data gained by experience.

The technical level of the thesis is adequate.

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

Assess correctness of usage of formal notation. Assess typographical and language arrangement of thesis.

The thesis is in most parts well structured and clearly written.

Selection of sources, citation correctness **B - very good.**

Present your opinion to student's activity when obtaining and using study materials for thesis creation. Characterize selection of sources. Assess that student used all relevant sources. Verify that all used elements are correctly distinguished from own results and thoughts. Assess that citation ethics has not been breached and that all bibliographic citations are complete and in accordance with citation convention and standards.

The references are adequate. Existing work is clearly distinguished from own results of the student.

ČVUT
ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

REVIEWER'S OPINION OF FINAL THESIS

Additional commentary and evaluation *Present your opinion to achieved primary goals of thesis, e.g. level of theoretical results, level and functionality of technical or software conception, publication performance, experimental dexterity etc.*
The presented thesis is mostly well written, clearly structured and technically correct. However, there remain a few open questions and issues. (1) Is the analysis of the feature distributions (section 4.3.1) used in the design of the classifier network? (2) The proposed network is essentially a binary classifier. Why are you using two outputs, softmax and cross entropy? Equivalently, you could use one output, sigmoid and binary cross entropy. (3) The presented learning curves (figure 4.6) show no overfitting and at the same time quite large loss values. Have you tried to use larger architectures? (4) What do you mean by "optimising the capabilities of a (trained) classifier"? (section 4.4.2) (5) It might seem that computing Shapley values for individual features will require re-training of the network for different feature combinations. How is this avoided? (6) Do you see a way how to formalise the task of model optimisation w.r.t. additional object states (here signal mass) which are available at training time but are not part of the classifier input?

III. OVERALL EVALUATION, QUESTIONS FOR DEFENSE, CLASSIFICATION SUGGESTION

Summarize thesis aspects that swayed your final evaluation. Please present apt questions which student should answer during defense.

The thesis presented by Ondřej Matoušek clearly and without any doubts fulfils the criteria of a bachelor graduation work.

I evaluate the submitted thesis with classification grade B - very good.

Date: **6.6.2023** Signature: Boris Flach