THESIS SUPERVISOR'S REPORT

I. IDENTIFICATION DATA

Thesis title:	**A New Trackster Linking Algorithm Based on Graph Neural Networks for the CMS Experiment at the Large Hadron Collider at CERN**
Author's name:	**Jekatěrina Jaroslavceva**
Type of thesis :	master
Faculty/Institute:	Faculty of Electrical Engineering (FEE)
Department:	Dept. of Cybernetics
Thesis reviewer:	Ondřej Chum
Reviewer's department:	Dept. of Cybernetics

II. EVALUATION OF INDIVIDUAL CRITERIA

Assignment **extraordinarily challenging**

How demanding was the assigned project?

Success in the project depends not only on the solution itself, which requires understanding of advanced machine learning techniques, but also requires basic understanding of physical principles underlying the particle collision phenomena.

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.

Jekaterina has delivered two methods, one based on a direct learnable classification of pairs of possibly linked trackleds, and one based on graph neural networks, both the approaches outperform currently used state of the art. The first method makes a decision based on the pair only, the latter takes into account graph neighborhood of a fixed depth. The method is replicable, it is implemented in the project repository, ready to be instantly used.

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.

The thesis was directly supervised by Dr. Felice Pantaleo, physicist in CERN. His evaluation letter is attached in "Additional commentary and evaluation" section. Here I follow his evaluation. Jekaterina has, in the middle of her internship in CERN, actively organized a department-wide presentation of her results. She reflected the comments and suggestions she received after the presentation in the final version of the thesis.

Technical level **A - excellent.**

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 technical level is high, all concepts are well explained.

Formal level and language level, scope of thesis **A - excellent.**

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 typeset in LaTeX, well organized and presented. It is written in very good English.

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?

Reference to related work and its selection is adequate. The contributions are clearly stated.

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.

Thesis evaluation by the local supervisor in CERN Dr. Felice Pantaleo follows:

Jekaterina's thesis delves into the complex physics of the upcoming High-Luminosity Large Hadron Collider (HL-LHC) upgrade and focuses on the development of innovative reconstruction algorithms to handle the increased data complexity. The thesis meticulously describes the physics principles underlying the research topic, providing a comprehensive understanding of the challenges involved. Additionally, Jekaterina offers a well-written introduction to graph neural networks, showcasing her ability to explain complex concepts in a clear and accessible manner.

One commendable aspect of Jekaterina's work is her meticulous study of various references to obtain a thorough understanding of the state-of-the-art techniques and models. This extensive literature review serves as a solid foundation for her research, enabling her to develop a model that addresses the fragmentation and reconstruction quality degradation of tracksters in the Compact Muon Solenoid (CMS) at LHC's new endcap High-Granularity Calorimeter (HGCAL). Jekaterina's dedication to staying informed about the latest advancements is evident in her comprehensive exploration of relevant references.

The abstract provided by Jekaterina offers valuable insights into the thesis's content and highlights its significance. Her investigation focuses on machine learning approaches, with a particular emphasis on Graph Neural Network (GNN) models, to enhance event reconstruction through improved calorimetric clustering. She proposes a trainable GNN-based algorithm for accumulating incomplete energy fragments into well-formed tracksters, which is integrated into the CMS Software package as a linking plug-in. Jekaterina's meticulous evaluation of the algorithm's clustering and physics reconstruction performance on simulation data demonstrates its superiority over the currently used rule-based state-of-the-art benchmark. Her thesis showcases the potential of the proposed model to enhance the CMS event reconstruction, even in the challenging environment of the HL-LHC.

Considering the comprehensive evaluation of Jekaterina's thesis, I wholeheartedly recommend it to be awarded a classification grade of A - excellent. This grade reflects the exceptional quality and significance of her work, including her detailed description of the underlying physics and the well-documented features of the developed model. Jekaterina's research demonstrates a commendable ability to integrate the latest advancements in machine learning, graph neural networks, and event reconstruction techniques.

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.

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

A difficult problem is addressed and efficiently solved, delivering the-state-of-the-art results. The student was initiative and working independently. The student researched various approaches of machine learning and applied them in a novel way to the problem. The thesis manuscript itself is well written and is interesting to read for both physicist (introducing machine learning approaches, such as graph neural networks), and machine learning community (describing physics processes and measurements in the collider in an accessible way and the application of machine learning approaches to processing of the measured data). Last but not least, a GPU-ready implementation is also a valuable output of the work. The thesis has all properties of an excellent thesis. As such it should be considered for the Dean's prize.

Date: **6.6.2023** Signature: