

## I. IDENTIFICATION DATA

<b>Thesis name:</b>	Machine learning for ttH mechanism Higgs boson detection from CERN ATLAS data
<b>Author's name:</b>	Jan Presperin
<b>Type of thesis :</b>	<input type="text"/>
<b>Faculty/Institute:</b>	<input type="text"/>
<b>Department:</b>	Dept. of Computer Science
<b>Thesis reviewer:</b>	doc. Boris Flach
<b>Reviewer's department:</b>	Dept. of Cybernetics

## II. EVALUATION OF INDIVIDUAL CRITERIA

<b>Assignment</b>	<input type="text"/>
<i>Evaluation of thesis difficulty of assignment.</i>	
The thesis task is adequate. It required from the student to gain partial knowledge in particle physics and of related measuring systems	

<b>Satisfaction of assignment</b>	<input type="text"/>
<i>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.</i>	
The thesis meets the assignment	

<b>Method of conception</b>	<input type="text"/>
<i>Assess that student has chosen correct approach or solution methods.</i>	
The chosen models and methods are adequate	

<b>Technical level</b>	<input type="text"/>
<i>Assess level of thesis specialty, use of knowledge gained by study and by expert literature, use of sources and data gained by experience.</i>	
The proposed methods are technically correct	

<b>Formal and language level, scope of thesis</b>	<input type="text"/>
<i>Assess correctness of usage of formal notation. Assess typographical and language arrangement of thesis.</i>	
The thesis is well structured and clearly written(with minor exceptions)	

<b>Selection of sources, citation correctness</b>	<input type="text"/>
<i>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.</i>	
The cited sources are adequate and the bibliography complies with standards	

<b>Additional commentary and evaluation</b>
<i>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.</i>
<b>Thesis strengths:</b>

- The thesis is well written and clearly structured. The presented methods are technically correct.
- The experimental setup and the amount of experiments done by the author is impressive. In particular, I find the setup for hyper-parameter search for each of the three approaches very convincing.

**Thesis weaknesses:**

- The presentation of the chosen evaluation criterion (p-value and significance) is in my view not concisely written. E.g. What is the H0 hypothesis in the considered task? How does it influences the predictor? It remains unclear whether the final goal is a multi-class predictor or a binary predictor (ttH against all other classes)

### III. OVERALL EVALUATION, QUESTIONS FOR DEFENSE, CLASSIFICATION SUGGESTION

*The thesis compared three machine learning approaches for classifying particle production events resulting from proton collisions in the Large Hadron Collider (CERN). The author has carefully chosen an adequate experimental setup including hyper-parameter search for each of the considered methods. The experimental evaluation and comparison of the methods is convincing. Methodically, there are in my view some remaining ambiguities w.r.t. to the chosen training objectives and final evaluation criteria.*

**Questions:**

- Assume you are given a multi-class predictor that provides posterior class probabilities for given inputs. The task is however to predict one of the classes against all others. How would you use the multi-class predictor for this task? How will the optimal strategy change, if the predictor is allowed to reject inputs given for classification?
- How would you justify the chosen decision strategy (4.2)? Is there a loss function that leads to it?

I evaluate handed thesis with classification grade

Date: 8.6.2022

Signature: Boris Flach