THESIS REVIEWER'S REPORT

I. IDENTIFICATION DATA

Thesis title:	Search for a charged Higgs boson decaying into a neutral Higgs boson and a W-boson Using MachineLearning with ATLAS Data
Author's name:	Afandizada Azad
Type of thesis :	master
Faculty/Institute:	Faculty of Electrical Engineering (FEE)
Department:	Computer Science
Thesis reviewer:	Daniel Novak
Reviewer's department:	Cybernetics

II. EVALUATION OF INDIVIDUAL CRITERIA

Assignment **challenging**

How demanding was the assigned project?

The project was challenging, integrating machine learning techniques with a complex physics analysis involving ATLAS data and simulation. It required interdisciplinary knowledge across particle physics, software development, and data science.

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 addressed all assigned tasks. He surveyed and optimized ML algorithms, applied them to simulated datasets, evaluated performance, investigated feature importance, and studied signal-background separation uncertainties.

Methodology **correct**

Comment on the correctness of the approach and/or the solution methods.

The use of neural networks with TensorFlow/Keras for classification, supported by SHAP and permutation analysis for feature importance, was highly appropriate. The basic ML methods were described.

Technical level **C - good.**

Is the thesis technically sound? How well did the student employ expertise in the field of his/her field of study? Does the student explain clearly what he/she has done?

The technical depth is sound, reflecting solid understanding of physics, ML, and statistical analysis. The student showed knowledge in interpreting ROC curves, training strategies, and expected limits with ATLAS data. I have several remarks:

1. Feature Engineering:
   - The selection of only 15 features from over 300 available is justified by previous work, but lacks a systematic evaluation. It would have strengthened the study to report performance changes when adding/removing features.
   - Some features such as MET or tau decay vertex were briefly mentioned but not deeply analyzed.

2.Validation and Overfitting:
   - The thesis identifies overfitting in low-mass scenarios (250 GeV), but the countermeasures are relatively limited (early stopping, weight rebalancing).
   - Cross-validation techniques or regularization approaches (dropout, L1/L2) are not discussed.

3.Uncertainty Quantification:
   - Uncertainties are acknowledged but only statistical ones are included. Systematic uncertainties, which are vital in ATLAS physics analyses should be also addressed.

Formal 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 mostly well-written, logically structured, and detailed. Some minor language issues exist (e.g., article usage, sentence clarity), but they do not significantly impact readability.

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 thesis includes appropriate citations from both scientific literature and ATLAS-related resources. The separation of original work and references is clear and properly attributed.

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.

While the introduction provides a good overview, a clearer discussion of the implications of detecting a charged Higgs boson and how this work fits into the broader ATLAS program would enhance impact.

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

The thesis presents a reasonable study at the intersection of machine learning and experimental particle physics. The candidate investigates the classification of charged Higgs boson events using simulated ATLAS data and ML models. The student demonstrates familiarity with both the physics context and the data analysis tools required.

The grade that I award for the thesis is **B - very good.**

Questions for the Student:

1. Feature Selection

Why did you limit the input features to 15? Did you consider an automatic feature selection method such as recursive feature elimination or Lasso regression?

2. Generalization and Overfitting

Given the high discrepancy between training and validation losses for the 250 GeV case, did you explore regularization techniques or deeper architectures to improve generalization?

3. Systematic Uncertainties

How would the inclusion of systematic uncertainties (e.g., jet energy scale, b-tagging efficiency) affect your results and the expected limits?

4. Alternative Classifiers

Why did you choose a fully connected neural network (MLP)? Did you experiment with or consider decision tree-based models such as XGBoost or ensemble methods?

5. Data Realism

Your study is based on simulated events. What challenges do you foresee when applying this approach to real ATLAS data?

Date: 4.6.2025 Signature: