

The Large Hadron Collider (LHC) [1] is a superconducting two-ring, protons and heavy ions collider installed in the 27 km-long LEP tunnel at CERN in Geneva. It provides pp collisions at an unprecedented center of mass energy  $\sqrt{s} = 13$  TeV. Four experiments are installed in the LHC interaction points to analyse the particles produced by the collisions in the accelerator. Each experiment is characterized by a peculiar design optimized on its specific physics program.

This thesis is performed within the ATLAS [2] experiment, which is designed to explore a wide range of physics topics, with the primary focus of improving our understanding of the fundamental constituents of matter and their interactions. Currently particle physics phenomenology is well described by the so called Standard Model (SM), a quantum field theory based on  $SU(2) \otimes U(1) \otimes SU(3)$  gauge symmetry. ATLAS is studying the processes predicted by the SM such as W,Z and top production and compares the measured cross sections with the model predictions. Events with electrons and photons in the final state are important signatures for many physics analyses envisaged at ATLAS: excellent performance in the electrons and photons reconstruction is essential to exploit the full physics potential of the detector, both in searches for new physics and in precision measurements. For instance, the good electron and photon reconstruction performance played a critical role in the discovery of the Higgs boson [3], announced by the ATLAS and CMS [4] Collaborations in 2012 and in the measurement of its properties.

Electrons and photons in ATLAS are reconstructed starting from energy deposits in the electromagnetic calorimeter and tracks from inner detector hits. An electron is defined as an object consisting of a cluster built from energy deposits in the calorimeter with a track pointing to it. A converted photon is a cluster matched to a conversion vertex, and an unconverted photon is a cluster matched to neither a track or a conversion vertex. The  $e$  and  $\gamma$  are reconstructed independently, so it is possible to reconstruct  $e$  or  $\gamma$  from the same clusters and tracks. After the  $e/\gamma$  are reconstructed, an ambiguity resolver algorithm is applied on them: if a particular object can be easily identified only as a photon (a cluster with no good track attached) or only as an electron (a cluster with a good track attached and no good photon conversion vertex), then only a photon or an electron object is stored for analysis; otherwise, both an electron and a photon object are created. At the reconstruction level only simple algorithms are used. They resolve only the simplest cases whereas many objects are flagged as ambiguous, leaving the final arbitration at the analysis level. The ambiguous objects classification can be approached with machine learning techniques, which provide better results with respect to simple cut-based selections. In particular in this thesis the usage of a supervised learning algorithm, called Gradient Boosted Decision Tree (GBDT) has been studied.

Typically the reconstruction algorithm provides both electron and photon candidates in  $\sim 8\%$  of the *True electrons* and  $\sim 31\%$  of the *True photons*. In this thesis three main scenarios are investigated. The first case is a classification of all doubly reconstructed objects, creating a model trained on them. It explores the possibility to replace the current classification algorithm with a gradient boosted decision tree for all electrons and photons candidates with no pre-classification from ambiguity resolver. This model achieves an AUC of  $\sim 0.9958$ . The *double reconstruction* model sets the theoretical

limit because this approach can not be used in standard data reconstruction due to the enormous memory cost of saving all objects as doubly reconstructed. To reduce double reconstructions to be saved, a loose resolution is applied: firstly the standard ambiguity (*old amb*) resolver used in Run 2 data taking was investigated, secondly a more optimized resolver (*new amb*) foreseen for the upcoming Run 3 is studied. Two new models are trained on top of the particles classified as ambiguous by these simple tools and they can also achieve excellent particle classification capabilities. In these cases the *old amb* model, which has been trained on objects flagged as “ambiguous” by the old ambiguity tool, achieves an AUC of  $\sim 0.9808$ , whereas the AUC of the *new amb* model, which is based on the new ambiguity tool, is equal to  $\sim 0.9907$ . The new ambiguity tool increases the performance of a GBDT based on ambiguous objects, allowing a better classification efficiency and a higher photon efficiency than the old one: if the photon efficiency  $\epsilon_{ph}$  is fixed the difference between the electron efficiencies  $\epsilon_{el}^{new} - \epsilon_{el}^{old}$  is in its maximum value equal to  $\sim 7\%$ ; on the contrary if the electron efficiency  $\epsilon_{el}$  is fixed the difference between the photon efficiencies  $\epsilon_{ph}^{new} - \epsilon_{ph}^{old}$  is in its maximum value equal to  $\sim 3\%$ .

As shown by these results, models trained on ambiguous objects flagged by ambiguity tools are able to achieve very good performance in classification. In particular, with the introduction of the new ambiguity tool, performance has improved since the ambiguous photons are increased allowing the BDT algorithm to reach the theoretical limit. Moreover, these two models can be used at the analysis level to classify ambiguous objects. In fact, the analyses are now flagging all the ambiguous objects as all electrons or all photons, while a GBDT provides a continuous score for them which allows for an optimal choice based on analysis needs.

# Bibliography

- [1] European organization for nuclear research, *LHC design report*, CERN libraries, Geneva (2004), <http://cds.cern.ch/record/782076/files/>.
- [2] The ATLAS Collaboration et al 2008 JINST3 S08003, <https://iopscience.iop.org/article/10.1088/1748-0221/3/08/S08003/pdf>
- [3] *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, The ATLAS Collaboration, In: *Phys. Lett. B* 716.arXiv:1207.7214.CERN-PH-EP-2012-218 (Aug. 2012) DOI:<https://doi.org/10.1016/j.physletb.2012.08.020>, URL:<http://www.sciencedirect.com/science/article/pii/S037026931200857X>
- [4] *The CMS experiment at the CERN LHC*, The CMS Collaboration <https://doi.org/10.1088/1748-0221/3/08/S08004>