The Large Hadron Collider (LHC) [**?**] is a superconducting two-ring, protons and heavy ions accelerator and collider installed in the 27 km-long LEP tunnel at CERN in Geneve (Ch). It provides pp collisions at an unprecedent center of mass energy $\sqrt{s}$ = 13 TeV. Inside the accelerator, two high-energy particle beams travel close to the speed of light before they are made to collide. Four experiments are installed in the LHC interaction points to analyse the myriad particles produced by collisions in the accelerator. Each experiment is characterized by a peculiar design optimazed on the specific physics program.

This thesis focuses on the ATLAS [**?**] 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 this is well described by the so called Standard Model (SM), a quantum field theory based on SU2⊗U1⊗SU3 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 a Higgs boson [3], announced by the ATLAS Collaboration in 2012 and in the measurement of its properties.

Electrons and photons in ATLAS are reconstructed starting from energy deposits in the EM calorimeter and tracks reconstructed in the inner detector. An electron is defined as an object consisting of a cluster built from energy deposits in the calorimeter and a matched track. A converted photon is a cluster matched to a conversion vertex (or vertices), and an unconverted photon is a cluster matched to neither an electron track nor a conversion vertex. After the objects are reconstructed, the ambiguity resolver is applied on them. Its purpose is that 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 created for analysis; otherwise, both an electron and a photon object are created.

At the reconstruction level only simple algorithms are used. They reconstruct only the simplest cases whereas the other 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 supervided learning algorithms, called Gradient Boosted Decision Tree (GBDT) have been studied.

Typically the reconstruction algorithm provides both electron and photon candidates in 10% of the *True electrons* and 30% of the *True photons*. The first case adressed in this thesis 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. This model

is able to correctly classify up to $\sim 99.58\%$ of particles in the *outer* test set. The *double reconstruction* model sets the theoretical limit because this approch can not used in standard data training due to the enormous memory cost of saving all objects as doubly reconstructed. To avoid this problem by reducing double reconstructions, a loose arbitration is applied.

The other two models can also achieve excellent particle classification capabilities: on the outer test set the *old amb* model, which has been trained on objects flagged as "ambiguous" by the old ambiguity tool, correctly classifies $\sim 98.08\%$ of particles, whereas the *new amb* model, which is based on the new ambiguity tool, $\sim 99.07\%$. The new ambiguity tool increases the performance of a GBDT based on ambiguous objects, allowing a better classification than the old one: if the photon efficiency ϵ_{ph} is fixed the difference between the electron efficiencies $\epsilon_{el}^{new} - \epsilon_{el}^{old}$ is in its maximum value equal to $\sim 3\%$; on the contrary if the electron efficiency ϵ_{el} is fixed the difference between the photon efficiencies $\epsilon_{ph}^{new} - \epsilon_{ph}^{old}$ is in its maximum value equal to $\sim 7\%$.

As shown by these results, models trained on ambiguous objects of ambiguity tools are able to achieve very good performance in classification approaching very close to the theoretical limit. In particular, with the introduction of the new ambiguity tool, performance has improved and therefore classification efficiency has improved as well.

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`