Track Reconstruction in High Energy Physics using Machine Learning

07/01/2019

**Topic:** Machine Learning and Deep Learning

The reconstruction of physical quantities of the particles stemming from collision events is a particularly important problem in modern High Energy Physics (HEP). For example, the purpose of the ATLAS experiment is to detect the products of a proton-proton collision inside the Large Hadron Collider (LHC) at CERN, Geneva. The raw data per event generated by LHC is around one million bytes (1 MB), produced at a rate of about 600 million events per second. The raw data recorded in a year by the ATLAS experiment is around 3.2 PB, and the amount of reconstructed data corresponds to nearly 2 PB. Moreover, further developments of the machine are going to increase these numbers considerably.

A crucial part of the data analysis is to determine the parameters of the vertices and particle trajectories for each event and to extract the tracks among the events that are defined as hits, which are produced by charged particles passing through a number of coordinated planes of tracking detectors. The track recognition therefore consists in joining these hits into clusters of the same track, while also discarding fake hits and noise.

A particularly interesting new way to tackle this problem is to use machine learning and deep learning. Deep learning techniques have achieved great results in pattern recognition tasks. These methods rely on useful representation for data, often using very low level inputs such as pixel in an image. In particular, neural networks are known to be very good at finding patterns and model non-linear dependencies in data. Convolution Neural Network have shown good results in rejecting ”seeds” (i.e. initial approximation of track parameters of charged particles) that do not correspond to real tracks by comparing the shapes of the hit clusters used in the seed. Neural network classifiers that determinate if the hit is caused by a single track or multiple tracks have been particularly efficient in dense regions. Finally using deep neural networks as classifier of fake hits (trajectories that don’t correspond to a real particle) has been proved to increase efficiency.
The aim of this project is to apply Recurrent Neural Networks (RNN) and Convoluted Neural Network (CNN) onto the data already gathered by ATLAS detectors, in particular the tracking system. The results obtained using these new methods will then be compared with the data analysed with the currently employed methods, which are inherently serial, in terms of detection efficiency and computing time.