



Exploration de méthode de machine learning pour la reconstruction d'événements en provenance de télescope Cherenkov avec CTA/LST-1

Spécialité Astrophysique

Niveau d'étude Bac+4/5

Formation Ingenieur/Master

Unité d'accueil [DPhP](#)

Candidature avant le 31/03/2024

Durée 4 mois

Poursuite possible en thèse non

Contact [de BONY de LAVERGNE Mathieu](#)
+33 1 69 083020
mathieu.debonyselavergne@cea.fr

Résumé

Le stage consiste en l'exploration de l'utilisation de différents algorithmes de machine learning, notamment les réseaux de neurones, dans la reconstruction d'événements avec CTA. Le projet vise à améliorer les performances de méthodes de reconstruction actuelles. Des connaissances en Python sont requises et en machine learning souhaitées.

Sujet détaillé

Les Imaging Atmospheric Cherenkov Telescopes (IACT) observent l'émission de rayons gamma de très haute énergie (VHE) provenant de sources astrophysiques. Pour accomplir cette tâche, le télescope détecte la lumière Cherenkov émise par les gerbes atmosphériques créées par les rayons gamma entrant dans l'atmosphère. Le LST-1 est un télescope prototype pour le Large Size Telescope (LST) de Cherenkov Telescope Array (CTA), la prochaine génération d'instruments qui étendra la gamme d'énergie et améliorera la sensibilité jusqu'à un ordre de grandeur par rapport aux expériences actuelles.

Un des défis de la technique d'observation IACT est de pouvoir, à partir de l'image prise par le télescope, déterminer la direction, l'énergie mais aussi le type de particules qui a créé la gerbe atmosphérique. Cette étape d'analyse est appelée reconstruction et est cruciale pour la performance de l'instrument. Une technique habituelle pour effectuer cela est d'utiliser la méthode dite de Hillas. Elle consiste à extraire des paramètres de l'image. Ces paramètres sont ensuite fournis à un algorithme de random forest afin de déterminer les paramètres de la particule qui a créé la gerbe.

La méthode Hillas associée aux random forest a démontré sa robustesse au fil du temps et est actuellement utilisée par plusieurs expériences comme MAGIC, VERITAS, et CTA. Mais l'utilisation des random forest a aussi quelques inconvénients, comme le surajustement, leur taille, etc. L'objectif de ce stage sera d'explorer d'autres algorithmes de machine learning comme les réseaux de neurones pour remplacer les randoms forest dans l'étape de reconstruction de l'analyse pour CTA/LST-1 et de comparer les performances physiques obtenues à la solution actuelle.

Le candidat doit être familier avec le langage Python. Quelques connaissances en data science et en machine learning seront un avantage.

Mots clés

CTA, Astronomie gamma, machine learning

Compétences

Python, Data science, Machine learning

Logiciels

Python, Pandas, Scikit-learn, Pytorch

Exploring machine learning for events reconstruction of Imaging Atmospheric Cherenkov Telescope data with the CTA/LST-1

Summary

Internship to explore machine learning, specifically neural networks, for event reconstruction with the CTA's LST-1. The project aims to enhance actual reconstruction performance. Proficiency in Python is required and in machine learning is desirable.

Full description

Imaging Atmospheric Cherenkov Telescopes (IACT) observe Very High Energy gamma-ray (VHE) emission from astrophysical sources. To be able to perform this task, the telescope detects the Cherenkov light emitted by the atmospheric showers created by the gamma-rays entering the atmosphere. LST-1 is a prototype telescope for the Large Size Telescope of Cherenkov Telescope Array, the next generation instrument of IACTs that will extend the energy range and improve the sensitivity up to one order of magnitude compared to current experiments.

One of the challenges of the IACT observation technique is to be able, from the image taken by the telescope, to determine the direction, the energy but also the type of particles that has created the atmospheric shower. This analysis step is called reconstruction and is crucial for the performance of the instrument. A usual technique to perform this is to use the so-called Hillas method. It consists in the extraction of parameters from the image. These parameters are then provided to a random forest algorithm in order to determine the parameters of the particle that has created the shower.

The Hillas method associated with random forests has shown its robustness across time and is currently used by several experiments like MAGIC, VERITAS, and CTA. But the use of random forests has also some drawbacks, like overfitting, their size, ... The objective of this internship will be to explore other machine learning algorithms like neural networks to replace the random forests in the reconstruction step of the analysis for CTA/LST-1 and compare the obtained physical performance to the current solution.

The candidate is expected to be familiar with the python language. Some knowledge about data science and machine learning will be an advantage.

Keywords

CTA, Gamma-ray Astronomy, machine learning

Skills

Python, Data science, Machine learning

Softwares

Python, Pandas, Scikit-learn, Pytorch