

Master 2 Internship

ANITI CertifAI – Towards the certification of ML-based systems / LAAS-CNRS

Title: Data generation for the evaluation of runtime monitoring mechanisms of neural networks

Keywords	Computer science, artificial intelligence, machine learning, data generation, safety monitoring
Technologies	Python, AI frameworks
Location	LAAS-CNRS, Toulouse
Level	Master 2
Duration	4–6 months
Profile	Computer science, artificial intelligence
Salary	~600€/month
Supervisors	Jérémie Guiochet, Mathieu Dario

Project description

Artificial intelligence (AI) is now used in an increasing number of critical applications, such as autonomous vehicles or surgical robotics. In these systems, perception tasks largely rely on deep neural networks (DNNs). However, their use remains limited because the behavior of these components is not guaranteed and raises reliability and safety issues. Runtime monitoring mechanisms are among the promising approaches to address these shortcomings. Their role is to continuously observe the neural network to prevent or detect a potential error. Two main families of monitors can be distinguished: rule-based monitors, which verify the satisfaction of safety properties; and data-driven monitors, which are themselves built through learning (often using neural networks). It is crucial to evaluate the performance of these monitors, particularly in situations where the main neural network may produce incorrect predictions.

This raises the question of what data should be used for such evaluation: how to collect it, how to generate it, and how to cover both nominal situations and so-called abnormal ones. The objective of the internship is to work on creating such data and to set up the framework to test one or more monitors provided by the AI research community.

The internship is part of a case study related to aviation, based on the LARD dataset (Landing Approach for Runway Detection). This corpus contains nearly 17,000 real and synthetic images representing different phases of aircraft approaches at several airports. It can also be enriched by generating new real or simulated data (using simulators such as Google Earth Studio). The internship will therefore focus on two main directions: the generation of realistic visual trajectories (nominal or likely to cause failures) using LAAS tools coupled with the LARD dataset; and the development of a test architecture to evaluate one or more monitors from the literature.

Application

To apply, please upload your **CV**, **cover letter** and **transcripts** of the last 2 years on the website <https://www.laas.fr/fr/offres/msc-internship-neural-network-monitoring/>