DEEL France-Quebec

COOPÉRATION FRANCE-QUÉBEC

NEWSLETTER #09

March 2021

TOULOUSE & QUEBEC JOIN FORCES TO DEVELOP AI FOR CRITICAL SYSTEMS

RLVS REINFORCEMENT LEARNING VIRTUAL SCHOOL

The **Reinforcement Learning Virtual School** (RLVS) takes place on a total of six days: March 25-26, April 1-2, and April 8-9, 2021.

This virtual school is an introduction to Reinforcement Learning aimed at masters and Ph.D. students, academics, and industrial researchers with a solid background in mathematics and computer science. It consists of lectures, hands-on sessions, and keynote talks from leading researchers in reinforcement learning and related fields.

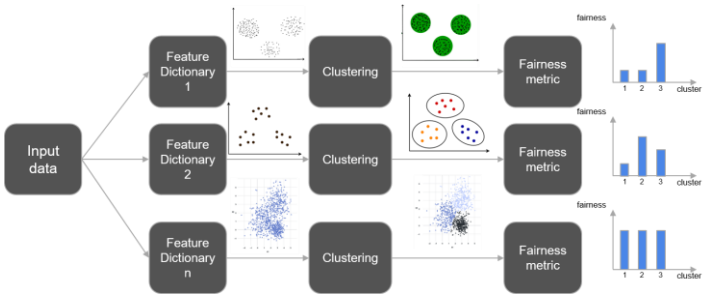
The school is free and entirely virtual. More details about how to access the event, the program, and the list of speakers, are available on our website: https://rlvs.aniti.fr/

This event is organized by ANITI with the support of Institut de Mathématiques de Toulouse - Université Toulouse III – Paul Sabatier, IRT Saint Exupéry, ISAE-SUPAERO, Laboratoire d'Analyse et d'Architecture des Systèmes – CNRS, Toulouse School of Economics - Université Toulouse I – Capitole, and Université Fédérale Toulouse Midi-Pyrénées. It is sponsored by DeepMind.

Replays are available → https://rl-vs.github.io/rlvs2021/

 # Sébastien GERCHINOVITZ (head organizer), David BERTOIN, Tom CESARI (Toulouse School of Economic), Nicolas MANSARD (LAAS/CNRS), Emmanuel RACHELSON (ISAE-SUPAERO)

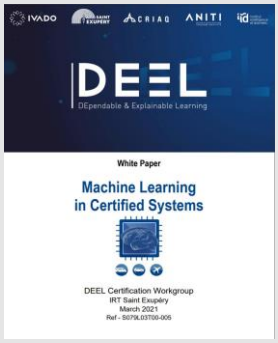
DETECTING BIAS IN IMAGE MODELS BY CLUSTERING



Learning a biased representation can be detrimental to the performance of predictive models at test time. As a means to tackle this problem, **we have developped a highly flexible heuristic tool to detect potential sources of bias and assess whether a trained model has learnt a biased representation**. It may thus be used as a stepping-stone to apply a bias correction technique without supervision by merging both techniques into a single pipeline.

We have performed an extensive analysis on **two DEEL industrial use-**cases (EuroSAT and Blinkers) and **one public dataset (CelebA)**, and with this information we have written a **comprehensive report** – available here for partners. **A tool has been developped jointly with example scripts and notebooks** to illustrate how to use it in different scenarii, with or without a trained model.

 # Agustin Martin PICARD, Jayant SEN GUPTA, Quentin VINCENOT



The excellent white paper about #machinelearning in certified systems performed by the DEEL Franco-Canadian project Dependable and Explainable Learning is available!

https://arxiv.org/abs/2103.10529

RESULTS OF THREE ACTIVITEES PRESENTED BY THE CERTIFICATION WORKGROUP

ACAS-XU: Approximating while compressing lookup tables (LUT) with a set of neural networks (NN) is an emerging trend in safety critical systems. We have explored how to compress **ACAS Xu LUT** while preserving the system safety and offering adequate means of certification. We have designed a proved hybrid architecture (mixing NN and a safety net), the assurance level of which comes from the quality of the learning and the use of formal methods. **An end to end certification strategy**, in line with on-going certification groups' activities, has been proposed to demonstrate the completeness and the correctness of the **ML-based system implementation with respect to the system and safety requirements (paper submitted to SafeComp'21)**

 # Mathieu DAMOUR, Florence DE GRANCEY, Christophe GABREAU, Adrien GAUFFRIAU, Jean-Brice GINESTET (DGA), Alexandre HERVIEU (DGA), Thomas HURAUX, Claire PAGETTI (ONERA), Ludovic PONSOLLE, Arthur CLAVIERE (COLLINS AEROSPACE)

COMPUTER VISION: The Computer Vision use case team provided ideas and methods to achieve trustworthiness of a perception application in the context of railways. The team developed a **dataset specification and verification strategy** (lien 3), **a proposition of standard for dataset definition** (lien 2) **and some recommendations for ML testing** (publication to come).
Lien 2 : https://arxiv.org/abs/2101.03020 & Lien 3 : https://arxiv.org/abs/2011.01799

 # Cyril CAPPI, Camille CHAPDELAINE, Laurent GARDES, Eric JENN, Baptiste LEFEVRE, Sylvaine PICARD, Thomas SOUMARMON

PROBABILISTIC ASSESSMENT: the "probabilistic assessment" team has continued is investigations of the relationship between the probabilistic approaches of safety risk and probabilities based analyses and properties of the machine learning. We interviewed some specialists of the safety of airborne and automotive electronics hardware, to understand their practices and rules of statistics and probabilities in the certification process. On the other side, we used an automotive use case to finely understand and evaluate how some mathematical properties (ie uncertainties bounding theorems) could serve to guarantee that the safety goal is reached. The exploration is vast, and the transfer of the assumptions to the reality, together with the guarantee of generalization remain hard points. So, a lot of tasks to be continued!

 # Jean-Luc ADAM, Lucian ALECU, Hugues BONNIN, Kevin DELMAS (ONERA), Jean-Marc GABRIEL, Sébastien GERCHINOVITZ, Franck MAMALET

DEEL QC DISCUSSION GROUP

Because the current pandemic context can make exchanges between project members difficult, and considering that synergy is an important issue for DEEL, essential to achieve the objectives of the project, we decided to add another recurrent virtual event to our agendas. We are glad to introduce the **DEEL Qc Discussion Group happening on the 3rd Wednesday of each month** at 4 pm (UTC+1) via Teams. The meetings last 1 hour (30 min of presentation and 30 min of discussion) and are an opportunity for students and other DEEL members to present their work regardless of their level of advancement.

The objective is to meet each other, promote exchanges and collaborations with the different partners of the project and bring an outside perspective to our research.

Hope to see many of you for our next edition on April 21st, which will deal with "Adversarial attacks in AI eXplanability

Le AN, Fanny EYBOULET, Yann PEQUINGOT

KEY DATES & INFORMATIONS

Certification Mission	Next workshops : 5th & 6th May – 26th & 27th May
Reinforcement learning School	Free and entirely virtual, total of 6 days : March 25th-26th, April 1st- 2sd and April 8th-9th
« Les Carrefours DEEL »	Find below the link of the last edition (1st April) https://youtube.com/playlist?list=PLe0GsO2LeMrKr9RXQhrLL9bEIdID0MMWr
MobiliT.AI 2021	Interactive and dynamic format : May 10th to 12th → Registration
Rendez-vous en Intelligence artificielle de Québec	Virtual edition : 12th & 13 th April → RIAQ Site