DEEL MAG #07

TOULOUSE & QUEBEC JOIN FORCES TO DEVELOP AI FOR CRITICAL SYSTEMS

FAIRNESS SEEN AS A GLOBAL SENSIVITY ANALYSIS

In 2013 Florida state gathered data in order to build an algorithm which would assess whether a criminal defendant is likely to be a recidivist. A few years later, a study discovered that, despite a first analysis showing that the predictions were not biased by the

Fairness definition	GSA measure associated
Statistical Parity	$\mathrm{Var}(\mathbb{E}[f(\mathbf{X})|S]) \to Sob_S(f(\mathbf{X}))$
Avoiding Disparate Treatment	$\mathbb{E}[\mathrm{Var}(f(\mathbf{X})|X)] \to SobT_S(f(\mathbf{X}))$
Equality of odds	$\mathbb{E}[\mathrm{Var}(\mathbb{E}[f(\mathbf{X})|S,Y]|Y)] \to CVM^{ind}(f(\mathbf{X}), S|Y)$
Avoiding Disparate Mistreatment	$\mathrm{Var}(\mathbb{E}[\ell(f(\mathbf{X}),Y)|S]) \to Sob_S(\ell(f(\mathbf{X}),Y))$

Table 2: Common fairness definitions and associated GSA measures

ethnic origin of the prisoners, the algorithm lacked efficiency for Afro-American prisoners. Actually Afro-American prisoners were ranked by the algorithm more severely than other categories, leading to discrimination in the odds. Hence one of the major task of research in the fairness for algorithmic decisions is the mere definition of the biased treatment of the machine, which implies being able to define and quantify fairness. Many different ways to do so exist, each one being specific to the learning setup.

In our work, inspired by the notion of risk analysis of critical systems and in particular the theoretical notion of sensitivity analysis, we define a generic mathematical framework that unifies any fairness measure used in supervised learning. This enables first to explain existing fairness metrics in an unified way, and second to propose new extensions and thus handle new measures for bias in industrial systems

 # Clément BENESSE, Thibaut BOISSIN, Jean-Michel LOUBES

WORKSHOP "MACHINE LEARNING IN CERTIFIED SYSTEMS"

The DEEL project is glad to announce the workshop "Machine Learning in Certified Systems: Theoretical and Practical Challenges" that will be held virtually on January 14th and 15th (afternoon), 2021. Various topics in mathematics, computer science, and industrial challenges, all related to certification of safety-critical systems with AI components, will be covered.

Registration is free but mandatory. All practical details, including the poster submission instructions, the list of invited speakers, and a full description of the program, can be found on our website:


https://mlcertifiedsystems.deel.ai/

 # Adrien GAUFFRIAU, Sébastien GERCHINOVITZ

KEY DATES & INFORMATIONS

Certification Mission	Next workshops : 27th & 28th January – 3rd & 4th Match – 31th & 1st April
Reinforcement learning School	Free and entirely virtual, total of 6 days : March 25th-26th, April 1st- 2sd and April 8th-9th
« Les Carrefours DEEL »	4th edition → 4th February
MobiliT.AI	Interactive and dynamic format : May 10th to 12th

PLATEAUS, LOCAL MINIMA, AND GLOBAL MINIMA OF DEEP LINEAR NETWORKS
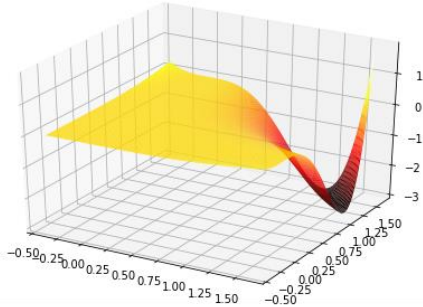
Neural networks are most often used with a number of parameters which is much larger than the number of examples in the training set. In this regime it is often observed that the optimization algorithm outputs parameters for which the network correctly fits the training set. Under which conditions this happens exactly is still a partially open question. Furthermore, in this regime, we could expect the network to generalize poorly to unseen data, which is not the case in many experiments. One possible explanation is that some regions of the parameter space are favored by the optimization algorithm; in other words, that some implicit regularization occurs.

To better understand these two phenomena, El Mehdi ACHOUR has studied **the landscape of the training error** as a function of the parameters. During the first year of his PhD, he has tightly described the locations and values of plateaus, local minima and global minima, which are regions where the optimization algorithms typically stop. These results for deep linear networks can be seen as a first step towards similar results for nonlinear deep networks.

 # El Mehdi ACHOUR, Sébastien GERCHINOVITZ, Francois MALGOUYRES

MAJOR DEEL EVENTS: "THE CARREFOUR DEEL & MOBILIT.AI "

The **"Carrefours DEEL"** are a transatlantic monthly seminar, organized in an interactive format in a virtual mode and specifically dedicated for the project partners and collaborators.

Work and results of the DEEL Project are presented by teams of the Institut de valorisation des données (**IVADO**), the **IRT Saint-Exupéry**, the Consortium de recherche et d'innovation en aérospatiale au Québec (**CRIAQ**), Artificial and Natural Intelligence Toulouse Institute (**ANITI**) and the Institut intelligence et données (**IID**) of Laval University.

On 3rd December took place the third edition of **the Carrefours DEEL**, on the subject of "**Explicability in machine learning**" presented by David VIGOUROUX and Thomas FEL. Video link available on demand by sending an email to equipe@iid.ulaval.ca.

The next events will take place on 2021, February the 4th.

2021
Coordinated by IVADO, IRT Saint-Exupéry, CRIAQ, and IID, we are pleased to announce : **the MobiliT.AI event** will now take place in 2021, May 10th to 12th in a very interactive and dynamic format.
Further information will be communicated later.

 # Ana GONZALEZ, Raymond POIRIER