Differential Privacy has Bounded Impact on Fairness in Classification

Paul Mangold*†[1], Michaël Perrot[1], Aurélien Bellet[1], and Marc Tommasi[1]

[1]Machine Learning in Information Networks – Inria Lille - Nord Europe, Centre de Recherche en Informatique, Signal et Automatique de Lille - UMR 9189 – France

Abstract

We theoretically study the impact of differential privacy on fairness in classification. We prove that, given a class of models, popular group fairness measures are pointwise Lipschitz-continuous with respect to the parameters of the model. We use this Lipschitz property to prove a high probability bound showing that, given enough examples, the fairness level of private models is close to the one of their non-private counterparts.

Keywords: Fairness, Differential Privacy, Classification, Machine Learning

*Speaker

†Corresponding author: paul.mangold@inria.fr