Machine learning as a tool to predict toxicity of chemicals

Marco Baity Jesi

Department of Systems Analysis, Integrated Assessment and Modelling, Eawag, Ueberlandstrasse 133,
CH-8600 Dübendorf, Switzerland
marco.baityjesi@eawag.ch

One of the pillars of modern civilization is the ability to synthesize and/or use an enormous range of chemicals, which allow for new and improved products and serve as pharmaceuticals, pesticides, food additives and the like. However, in order to maintain a healthy biodiversity, it is crucial to know the impact of these chemicals on the organisms of the ecosystems where they are diffused. Usually this is done through in-vivo testing. Yet, beyond its obvious ethical implications, in-vivo testing is not scalable to large amounts of taxa and chemicals, because of its strong requirements in terms of time, money, and highly-trained personnel. Machine learning comes in as a viable alternative with the potential of allowing us to explore the impact of large numbers of chemicals on many taxa: harmlessly, quickly and cheaply. We develop machine learning models that can quickly infer the mortality of a chemical on a certain fish species and highlight the importance of including information on both chemical and taxonomy as input features. We then analyze ways of comparing the reproducibility of in-vivo experiments with machine learning models trained on similar data, finding that there is no perfect method. Although comparisons between in-vivo and in-silico methods need to be taken with a grain of salt, machine learning presents as a potentially valid alternative to animal testing.