Predictive Deep Neural Network Models for ADME-Tox Properties: Learning from Large and Small Data Sets

Jan Wenzel,[a] Hans Matter,[b] Friedemann Schmidt[a]

Sanofi-Aventis Deutschland GmbH, [a]R&D, Preclinical Safety, Global Operations Germany and [b]R&D, Integrated Drug Discovery, Industriepark Höchst, 65926 Frankfurt am Main, Germany
E-mail: jan.wenzel@sanofi.com

While publically available contents and chemotype coverage of ADME-Tox databases are constantly growing, deep neural nets (DNN) emerged as transformative technology to analyze such large datasets and to predict ADME properties and even drug adverse effects. Particular ADME-Tox datasets are often sparsely populated, but specifically DNNs allow to overcome some limitation of classical machine learning; MT-DNN can ultimately combine different endpoints in a predictive multitask network. [1]

In this talk, we will describe a fully industrialized approach to parametrize and optimize the setup, training, application, and visual interpretation of DNNs to model ADME-Tox data. Quantitative regression models (>50k data) targeting microsomal metabolic lability (ML), passive permeability, and lipophilicity will be described. [1] Focusing on compound safety classification endpoints, e.g. phototoxcicity, it can be shown that DNNs based on comparably smaller data sets (< 5k) can also exhibit strong predictive power. On most studied datasets, MT-DNNs of multiple ADME-Tox properties perform statistical superior in comparison to DNN single-task models.

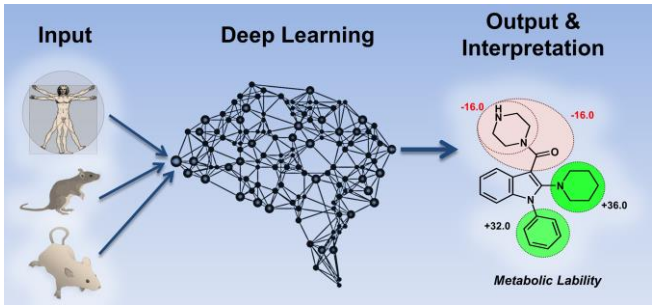

***Figure 1**: Microsomal metabolic lability data from human, rat and mouse are combined in a single predictive MT-DNN, which improves the external predictive power from R_{ext}^2 of 0.6 (single task) to R_{ext}^2 0.7 (MT-DNN). For interpretation and visualization, response maps are used to detect local property gradients based on structure fragmentation and derivatization. [1]*

[1] Wenzel J, *et al.*, *J Chem Inf Model*. **2019;** doi:10.1021/acs.jcim.8b00785