**Explainable AI for Pharmacophore-Based Drug Activity Prediction**

Joanna Ceklarz1, Krystyna Waniová 2, Wim Dehaen1, Tomasz Danel3, Martin Šícho1

*1 CZ-OPENSCREEN: National Infrastructure for Chemical Biology, Department of Informatics and Chemistry, Faculty of Chemical Technology, University of Chemistry and Technology Prague, Technická 5, 166 28 Prague, Czech Republic*
*2 Faculty of Mathematics and Computer Science, Jagiellonian University, Kraków, Poland*

*3 Faculty of Chemistry, Jagiellonian University, Kraków, Poland*

Pharmacophore representations are commonly used by medicinal chemists to identify and visualize structures necessary for biological function. Using them as representations of molecules for Graph Neural Network (GNN) training has yet untapped potential in deep learning. Deep learning architectures, to which GNNs belong, despite their excellent performance in many areas, come with one critical disadvantage - reduced or non-existent comprehensibility on *how* they reach their results. Many GNN-specific methods have been developed to answer questions about both feature and structural importance. When combined with the proposed pharmacophore representations, these methods could provide valuable insights to model users. Their application to chemical data, however, remains largely unexplored. We have developed two GNN models, a Graph Convolutional Network and a Graph Isomorphism Network, trained on 2D pharmacophore representations of small molecules, for drug activity prediction. We compare the results against shallow models, and against GNNs trained on traditionally used atomic representations of molecules. Using a selection of techniques, we aim to explain results of such models. We plan to obtain both local (molecule-level) and global (model-level) explanations, allowing us to analyse individual predictions as well as overarching model behaviour, to help identify the sources of errors and refine our models accordingly.