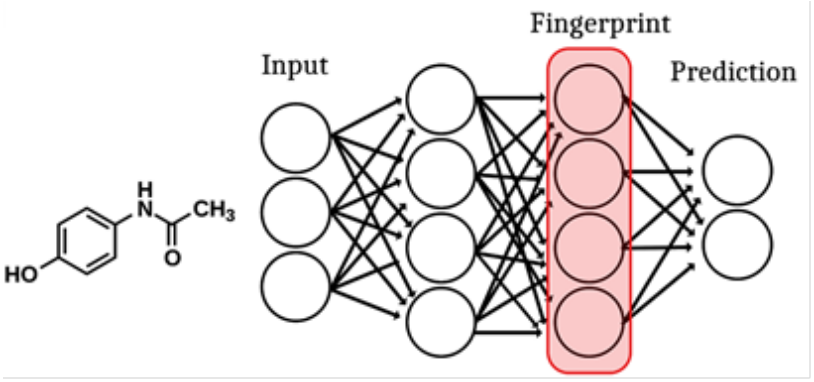
Neural Fingerprints: Structure- and activity-sensitive molecular representations based on neural networks for virtual screening approaches

Janosch Menke, **Oliver Koch**

Institute of Pharmaceutical and Medicinal Chemistry, University of Münster

Similarity-based virtual screening remains an important technique in the early stages of the drug discovery process. Amongst other things, the success relies on the appropriate choice of the underlying molecular representation, the molecular fingerprint. Our work focuses on improving these molecular representations to encapsulate more domain-relevant information with the help of neural networks. This approach works by extracting activations of the last hidden layer of a trained neural network as a novel neural network fingerprint representation for similarity-based virtual screening.



We could show for kinase inhibitors [1] and natural products [2] that the neural fingerprints extracted from trained neural networks outperform other fingerprints in similarity search, by providing overall more active hits than any other. So, it is possible to generate domain-specific neural fingerprints as a structure- or activity-sensitive molecular representation through the usage of supervised training for neural networks. Interestingly, we found that GNNs, compared to simple MLPs, created worse neural fingerprints when trained on the same tasks. Additionally, we were able to extract a Natural Product Likeness Score[2], as an alternative measure of assessing how likely a molecule is a natural product.

References

1. J. Menke, O. Koch, Using Domain-Specific Fingerprints Generated Through Neural Networks to Enhance Ligand-Based Virtual Screening. *J. Chem. Inf. Model.,* **2021**, *61*, 664-675.
2. J. Menke, J. Massa, O. Koch, Natural product scores and fingerprints extracted from artificial neural networks. *Comput. Struct. Biotechnol. J.,* **2021**, *19*, 4593-4602.