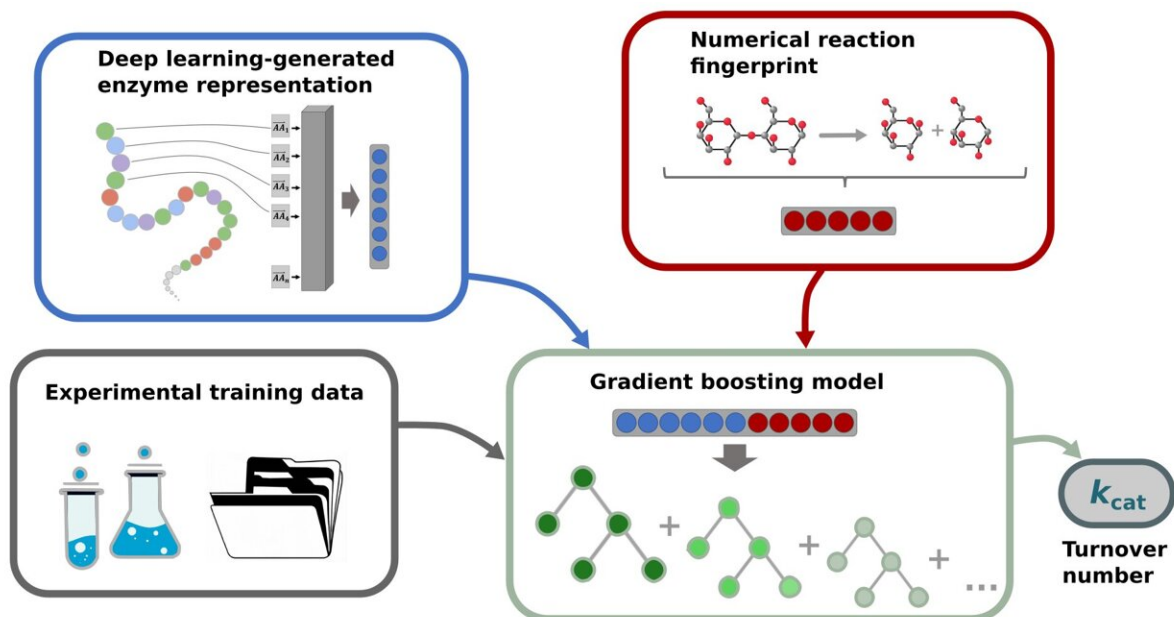# AI tool predicts the work rate of enzymes

July 24 2023

Schematic diagram of the prediction process for the turnover numbers of enzymatic reactions: Enzymes are amino acid sequences; these sequences are converted into numerical vectors, depicted as grey squares, which are then transformed into a single vector by a deep learning model (top left). Information about the catalysed reactions is also converted into numerical vectors (top right). Experimentally determined turnover numbers (bottom left) are used to train a gradient boosting model in order to predict the kcat turnover number (bottom right). Gradient boosting models are an ensemble of multiple decision trees, depicted in different green tones. Credit: HHU/Alexander Kroll

Enzymes play a key role in cellular metabolic processes. To enable the

quantitative assessment of these processes, researchers need to know the so-called "turnover number" (for short: k_{cat}) of the enzymes. In the journal *Nature Communications*, a team of bioinformaticians from Heinrich Heine University Düsseldorf (HHU) now describes a tool for predicting this parameter for various enzymes using AI methods.

Enzymes are important biocatalysts in all living cells. They are normally large proteins, which bind smaller molecules—so-called substrates—and then convert them into other molecules, the "products."

Without enzymes, the reaction that converts the substrates into the products could not take place, or could only do so at a very low rate. Most organisms possess thousands of different enzymes. Enzymes have many applications in a wide range of biotechnological processes and in everyday life—from the proving of bread dough to detergents.

The [maximum speed](#) at which a specific enzyme can convert its substrates into products is determined by the so-called turnover number k_{cat}. It is an important parameter for quantitative research on enzyme activities and plays a key role in understanding cellular metabolism.

However, it is time-consuming and expensive to determine k_{cat} turnover numbers in experiments, which is why they are not known for the vast majority of reactions. The Computational Cell Biology research group at HHU headed by Professor Dr. Martin Lercher has now developed a new tool called TurNuP to predict the k_{cat} turnover numbers of enzymes using AI methods.

To train a k_{cat} prediction model, information about the enzymes and catalyzed reactions was converted into numerical vectors using [deep learning models](#). These numerical vectors served as the input for a [machine learning model](#)—a so-called gradient boosting model—which predicts the k_{cat} turnover numbers.

Lead author Alexander Kroll said, "TurNuP outperforms previous models and can even be used successfully for enzymes that have only a low similarity to those in the training dataset." Previous models have not been able to make any meaningful predictions unless at least 40% of the enzyme sequence is identical to at least one enzyme in the training set. By contrast, TurNuP can already make meaningful predictions for enzymes with a maximum sequence identity of 0 to 40%.

Professor Lercher adds, "In our study, we show that the predictions made by TurNuP can be used to predict the concentrations of enzymes in living cells much more accurately than has been the case to date."

In order to make the prediction model easily accessible to as many users as possible, the HHU team has developed a user-friendly web server, which other researchers can use to predict the k_{cat} turnover numbers of enzymes.

More information: Alexander Kroll et al, Turnover number predictions for kinetically uncharacterized enzymes using machine and deep learning, *Nature Communications* (2023). DOI: 10.1038/s41467-023-39840-4

Provided by Heinrich-Heine-Universität Düsseldorf