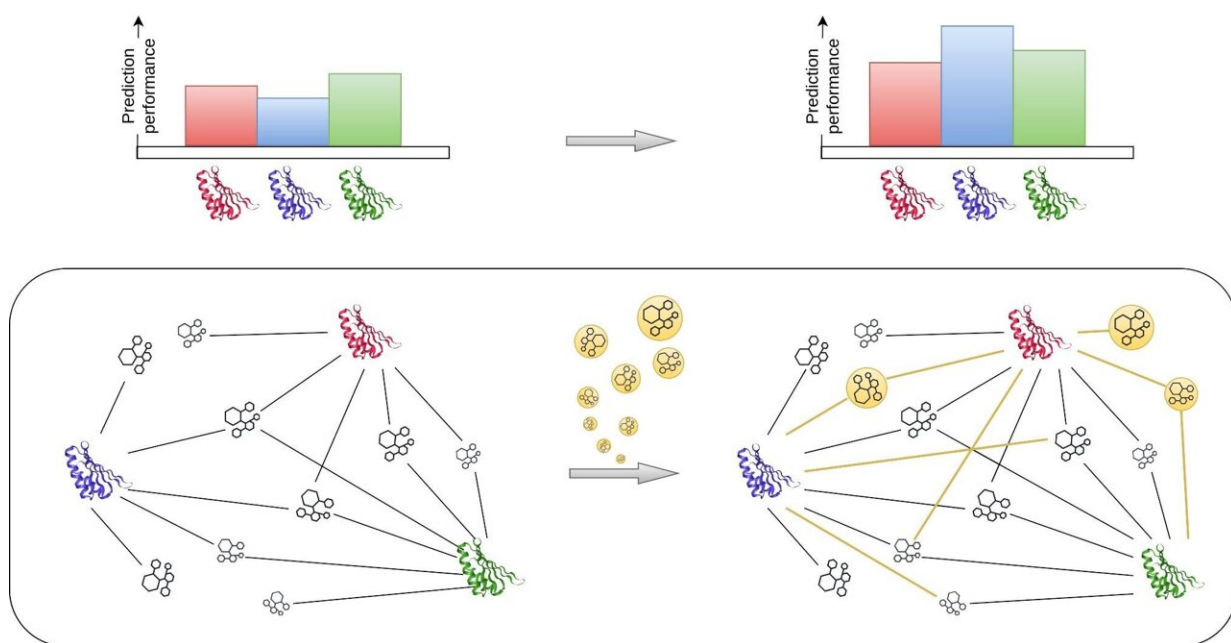


Advancing drug discovery through multitask learning techniques

March 28 2023



Effect of data enrichment on multitask learning. Credit: Ekaterina Sosnina/Skoltech

A team of researchers from the Skolkovo Institute of Science and Technology, the University of Vienna, and Sirius University of Science and Technology has published a study in the *Journal of Computer-Aided Molecular Design* that presents guidelines for enhancing the drug discovery process by utilizing multitask learning techniques.

In universities, students often take related courses, such as physics and mathematics, which leads to a better understanding of both subjects. Similarly, learning a new language is easier for those who have already had experience with languages, particularly cognate ones. The same principle applies to [machine learning](#), where a neural network can better comprehend multiple "subjects" if it learns them simultaneously. Considering that neural networks are one of the best approaches for predicting the biological properties of new chemical compounds, the question arises: How can we assist a [neural network](#) in simultaneously learning and predicting the properties of chemical compounds in relation to multiple biological targets?

Researchers analyzed three [datasets](#) for this purpose: a dataset with information on the antiviral activity of molecules and two datasets with information on the impact of molecules on various proteins in our body. The datasets varied in the completeness of information on each protein or virus. During the study, scientists discovered that adding data to the dataset is an effective method for improving prediction accuracy. Furthermore, they demonstrated that the more informative the original dataset is, the more noticeable this improvement. As a result of this work, the research team prepared a set of recommendations for using data enrichment technology to improve the quality and stability of predictions, as well as methods for objectively evaluating the improvement achieved.

"Multitask learning is widely used in many scientific fields. Unsurprisingly, it is increasingly being applied to develop new drugs. However, the potential of this approach has not yet been fully explored, presenting us with numerous unresolved tasks," the lead author of the study, Skoltech Ph.D. candidate Ekaterina Sosnina, notes. "We were inspired by the possibility of using multitask learning to develop new drug candidates and searched for ways to improve this approach. Following our recommendations, researchers in [drug discovery](#) will

enhance the predictive accuracy of their models and accelerate the identification of novel drug candidates."

More information: Ekaterina A. Sosnina et al, Improvement of multi-task learning by data enrichment: application for drug discovery, *Journal of Computer-Aided Molecular Design* (2023). [DOI: 10.1007/s10822-023-00500-w](https://doi.org/10.1007/s10822-023-00500-w)

Provided by Skolkovo Institute of Science and Technology

Citation: Advancing drug discovery through multitask learning techniques (2023, March 28) retrieved 9 April 2024 from <https://phys.org/news/2023-03-advancing-drug-discovery-multitask-techniques.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--