

## Purdue creates new framework to develop better drugs

01 January 2020 | News

**The framework, called Lemon, helps drug researchers better mine the Protein Data Base (PDB)**



US based Purdue University drug discovery researchers have created a new framework for mining data for training machine learning models.

The framework, called Lemon, helps drug researchers better mine the Protein Data Base (PDB) – a comprehensive resource with more than 140,000 biomolecular structures and with new ones being released every week.

The Lemon software platform is a fast C++11 library with Python bindings that mines the PDB within minutes. Loading all traditional mmCIF files in the PDB takes about 290 minutes, but Lemon does this in about six minutes when applying a simple workflow on an 8-core machine.

Lemon allows the user to write custom functions, include it as part of their software suite, and develop custom functions in a standard manner to generate unique benchmarking datasets for the entire scientific community.