

Large Scale Imputation of Drug Discovery Data using Deep Learning

Written by Benedict Irwin

Wednesday, 27 May 2020 09:14 - Last Updated Wednesday, 27 May 2020 09:48

Held on 26 May 2020.

This webinar was presented by Ben Irwin and Matthew Segall (Optibrium), Tom Whitehead (Intellegens) and Scott Rowland (Takeda Pharmaceutical).

We presented a case study in which Alchemite was applied to a data set comprising approximately 700,000 compounds and 1,000 experimental endpoints. This explored a variety of applications, including imputation and prediction of compound activities in project contexts, high-throughput screening results and a diverse range of ADME properties.

Unlike conventional machine learning approaches, Alchemite learns directly from sparse and noisy experimental data across multiple endpoints, typical of those available in drug discovery. In combination with molecular descriptors, this enables it to learn directly both from correlations between experimental endpoints as well as structure-activity relationships to make better predictions. Furthermore, the model provides a robust estimate of the confidence in each prediction, enabling attention to be focused on only the most accurate results.

To hear the narration, please increase your speaker volume.