Understanding the Sources of Performance in Deep Drug Response Models Reveals Insights and Improvements

Nikhil Branson
Digital Environment Research Institute
Queen Mary University of London

Pedro R. Cutillas
Barts Cancer Institute
Queen Mary University of London

Conrad Bessant
Digital Environment Research Institute
Queen Mary University of London
c.bessant@qmul.ac.uk

Abstract

Anti-cancer drug response prediction (DRP) using cancer cell lines plays a vital role in stratified medicine and drug discovery. Recently there has been a surge of new deep learning (DL) models for DRP that improve on the performance of their predecessors. However, different models use different input data types and neural network architectures making it hard to find the source of these improvements. Here we consider multiple published DRP models that report state-of-the-art performance in predicting continuous drug response values. These models take the chemical structures of drugs and omics profiles of cell lines as input. By experimenting with these models and comparing with our own simple benchmarks we show that no performance comes from drug features, instead, performance is due to the transcriptomics cell line profiles. Furthermore, we show that, depending on the testing type, much of the current reported performance is a property of the training target values. To address these limitations we create novel models (BinaryET and BinaryCB) that predict binary drug response values, guided by the hypothesis that this reduces the noise in the drug efficacy data. Thus, better aligning them with biochemistry that can be learnt from the input data. BinaryCB leverages a chemical foundation model, while BinaryET is trained from scratch using a transformer-type model. We show that these models learn useful chemical drug features, which is the first time this has been demonstrated for multiple DRP testing types to our knowledge. We further show binarising the drug response values is what causes the models to learn useful chemical drug features. We also show that BinaryET improves performance over BinaryCB, and over the published models that report state-of-the-art performance. See full paper at https://www.biorxiv.org/content/10.1101/2024.06.05.597337v2.