Towards a more inductive world for drug repurposing approaches

Jesus de la Fuente Cedeño (/profile?id=~Jesus_de_la_Fuente_Cede%C3%B1o1),
Guillermo Serrano (/profile?email=gserranos%40unav.es),
Uxía Veleiro (/profile?id=~Ux%C3%ADa_Veleiro1),
Mikel Casals (/profile?id=~Mikel_Casals1),
Laura Vera (/profile?email=lvera6%40unav.es),
Marija Pizurica (/profile?email=mpizuric%40stanford.edu),
Antonio Pineda-Lucena (/profile?email=apinedal%40unav.es),
Idoia Ochoa (/profile?id=~Idoia_Ochoa1),
Silve Vicent (/profile?email=silvevicent%40unav.es),
Olivier Gevaert (/profile?id=~Olivier_Gevaert1),
Mikel Hernaez (/profile?id=~Mikel_Hernaez1)

Published: 25 Oct 2023, Last Modified: 10 Dec 2023 AI4D3 2023 Oral Everyone Revisions (/revisions?id=ZhBwSxbJbu) BibTeX

Keywords: Drug discovery, Learning on graphs, Benchmarking datasets, Biologically driven benchmarking metrics, In-vitro validation

TL;DR: This study evaluates DTI prediction models, proposing and validating novel techniques for future design of drug repurposing approaches.

Abstract:
Drug-target interaction (DTI) prediction is a challenging, albeit essential task in drug repurposing. Learning on graph models have drawn special attention as they can significantly reduce drug repurposing costs and time commitment. However, many current approaches require high-demanding additional information besides DTIs that complicates their evaluation process and usability. Additionally, structural differences in the learning architecture of current models hinder their fair benchmarking. In this work, we first perform an in-depth evaluation of current DTI datasets and prediction models through a robust benchmarking process, and show that DTI prediction methods based on transductive models lack generalization and lead to inflated performance when evaluated as previously done in the literature, hence not being suited for drug repurposing approaches. We then propose a novel biologically-driven strategy for negative edge subsampling and show through in vitro validation that newly discovered interactions are indeed true. We envision this work as the underpinning for future fair benchmarking and robust model design. All generated resources and tools are publicly available as a python package.

Submission Number: 11