ARIA

APPLIED RESEARCH IN ACTION

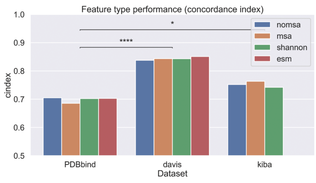
GNNs with Protein Dynamics for Enhanced Drug Targeting

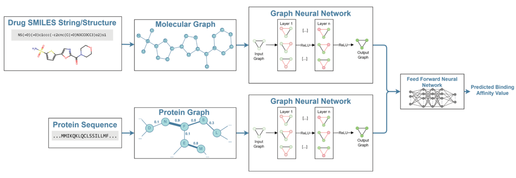
Analysis of binding affinity prediction models using Graph Neural Networks with modified protein features.

Jean Charle Yaacoub

Arvind Gupta, Huaxiong Huang and Faisal Habib
ACADEMIC SUPERVISORS

Sushant Kumar
INDUSTRY SUPERVISOR





PROJECT SUMMARY

The distinct mutations found in cancer-driving genes can influence the interaction between drugs and their intended targets/proteins, known as the drug-target binding affinity (DTA). Consequently, the efficacy of drug-based interventions is altered. To address this issue, the creation of dependable and precise models for predicting DTA becomes pivotal in optimizing drug treatment strategies. Recent advancements have employed graph-based representations for both drug and target proteins for DTA prediction. However, existing methodologies tend to overlook the dynamic nature of proteins (relegating it to a trainable aspect). In our study, we delved into the implications of incorporating protein dynamics within the protein graph representation, and we found that alterations to protein features made slight but no significant improvements to model performance. We also assess the applicability of our model by evaluating on the PDBbind dataset and the smaller PLATINUM dataset. These assessments revealed that the DTA prediction task is more challenging than initially presented and considerations like protein overlap in the training and test sets must be made in order to accurately gauge the model's robustness against mutational differences, indicating its potential utility for precision oncology.

REFERENCES

Jiang, M., Li, Z., Zhang, S. et al. (2020). Drug-target affinity prediction using graph neural network and contact maps. RSC advances 10, 20701-20712

Davis, M., Hunt, J., Herrgard, S. et al. (2011). Comprehensive analysis of kinase inhibitor selectivity. Nat Biotechnol 29, 1046–1051. https://doi.org/10.1038/nbt.1990

Tang, J., Szwajda, A., Shakyawar, S., et al. (2014). Making sense of large-scale kinase inhibitor bioactivity data sets: a comparative and integrative analysis. Journal of chemical information and modeling, 54(3), 735–743. https://doi.org/10.1021/ci400709d

Liu Z., Su M., Han L., et al. (2017). Forging the Basis for Developing Protein-Ligand Interaction Scoring Functions. Acc Chem Res. 50(2):302-309. https://doi.org/10.1021/acs.accounts.6b00491

Pires, D. E., Blundell, T. L., & Ascher, D. B. (2015). Platinum: a database of experimentally measured effects of mutations on structurally defined protein-ligand complexes. Nucleic acids research, 43(Database issue), D387–D391. https://doi.org/10.1093/nar/gku966

Jumper, J., Evans, R., Pritzel, A. et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589. https://doi.org/10.1038/s41586-021-03819-2

Lin, Zeming, et al. (2022). Evolutionary-scale prediction of atomic level protein structure with a language model. bioRxiv 2022.07.20.500902; doi: https://doi.org/10.1101/2022.07.20.500902

Diego del Alamo, Davide Sala, Hassane S Mchaourab, Jens Meiler (2022) Sampling alternative conformational states of transporters and receptors with AlphaFold2. eLife 11:e75751. https://doi.org/10.7554/eLife.75751

UHN Princess Margaret Cancer Centre

Computer Science UNIVERSITY OF TORONTO | Master of Science in Applied Computing