[BIOINFO 2024]

Cross-Attention Mechanisms for Drug Repositioning via Multi-modal Frameworks

Jae-woo Chu[1], Jong-Hoon Park[1], and Young-Rae Cho[2,3,*]

[1]*Computer Science Graduate School, Yonsei University*
[2]*Division of Software, Yonsei University - Mirae Campus*
[3]*Division of Digital Healthcare, Yonsei University - Mirae Campus*
**Corresponding author: youngcho@yonsei.ac.kr*

The adoption of artificial intelligence in new drug development is increasingly prevalent. It is particularly used for predicting drug properties and interactions through various modalities. This study proposes 'CARM', a cross attention for drug repositioning based multi-modal frameworks, for more precise drug-disease association predictions as part of drug repositioning techniques. CARM integrates drug images, graphs, molecular fingerprints, and disease similarity modalities, employing multi-head self-attention and cross-attention mechanisms for learning. The encoder, equipped with self-attention, learns relationships within each modality, while the decoder, utilizing cross-attention, learns interactions between different modalities before integrating them. The integrated feature vector is transformed into prediction probabilities through convolutional and fully connected layers to predict the correlations between drugs and diseases. The data used in the experiments were collected from DrugBank for drugs, OMIM for diseases, and Cdataset for drug-disease associations, using 656 drugs, 285 diseases, and 1,718 associations, excluding three drugs that could not be converted into molecular fingerprints.

Various experiments validating CARM's performance showed up to a 29.5% improvement in ROC-AUC compared to prior drug-disease association models using the CDataset. Additionally, the first ablation study on modality feature extraction investigated the optimal methods for each modality. It demonstrated that employing CNN for images, GIN and GAT for graphs, and FCFP along with Pharm2D for molecular fingerprints led to an increase in performance by up to 11.2%. A second ablation study comparing the use of single and multiple modalities demonstrated that integrating three modalities yielded up to an 8.9% performance improvement. Lastly, a generalizability assessment using additional benchmark datasets, Gdataset and LRSSL, also showed high generalizability with a ROC-AUC of over 0.8 across various datasets.