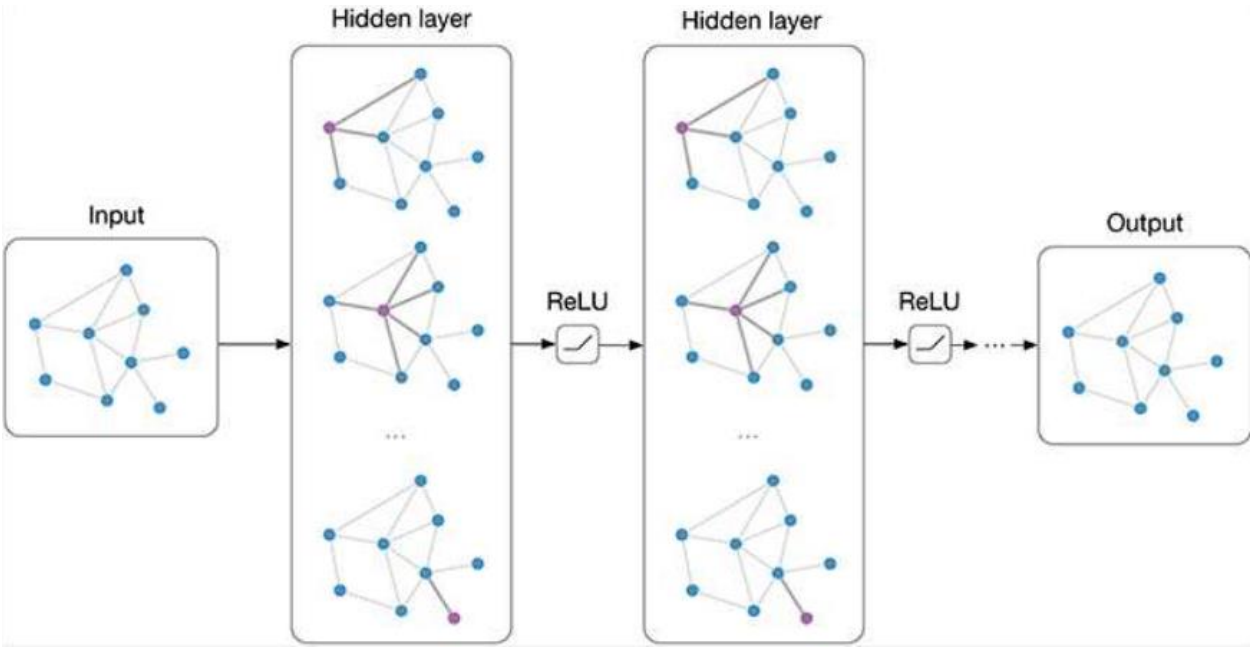
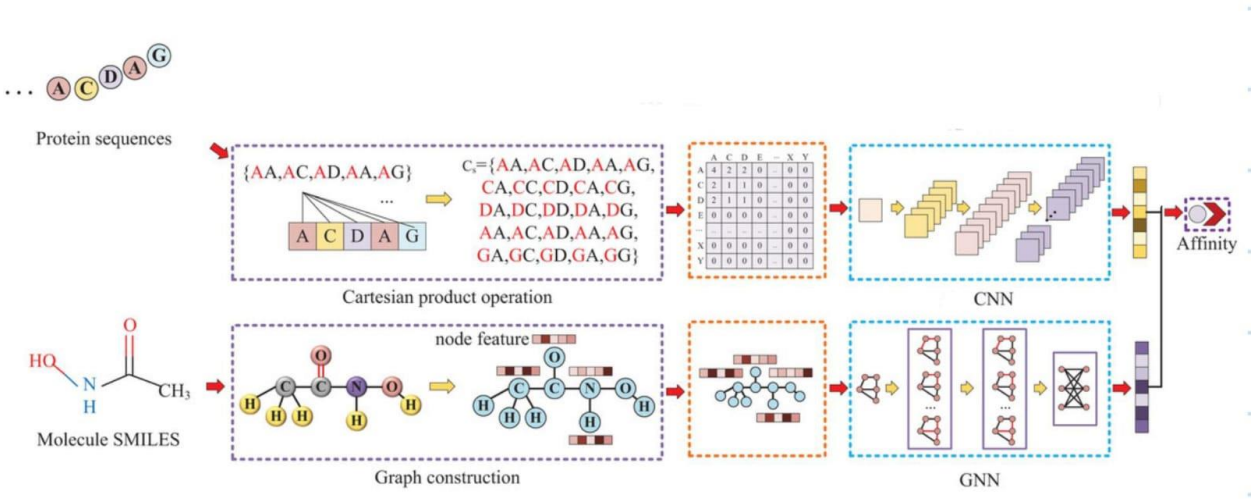
Drug–Target Interaction Prediction using GIN and CNN

Theoretical Background: Graph Isomorphism Network (GIN) Graph Isomorphism Network (GIN) is a type of graph neural network that achieves the expressive power of the Weisfeiler-Lehman test, a strong heuristic for graph isomorphism. GIN updates node features by aggregating features from neighboring nodes followed by an MLP, offering superior capability in distinguishing different graph structures compared to other GNN models like GCN and GAT.

Binding Affinity and Dissociation Constant (K_d) Binding affinity measures how strongly a drug molecule binds to its target protein. The dissociation constant (K_d) is used as an indicator—the lower the K_d, the stronger the binding. Predicting K_d accurately helps identify effective drug candidates and eliminate weak binders early in the drug discovery pipeline.

Methodology: GIN + CNN for Drug–Target Prediction Our model uses GIN to encode molecular graphs generated from SMILES representations and a CNN to process protein sequences. Features from both networks are concatenated and passed to a fully connected layer for affinity prediction. Dataset used : ***DAVIS***

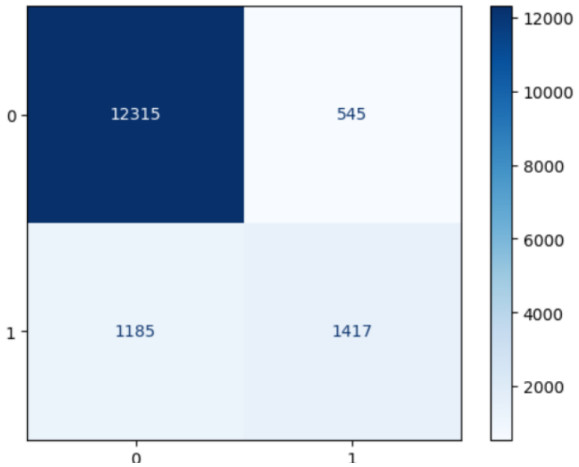




Evaluation and Results We use confusion matrices to evaluate model performance on training and testing splits.

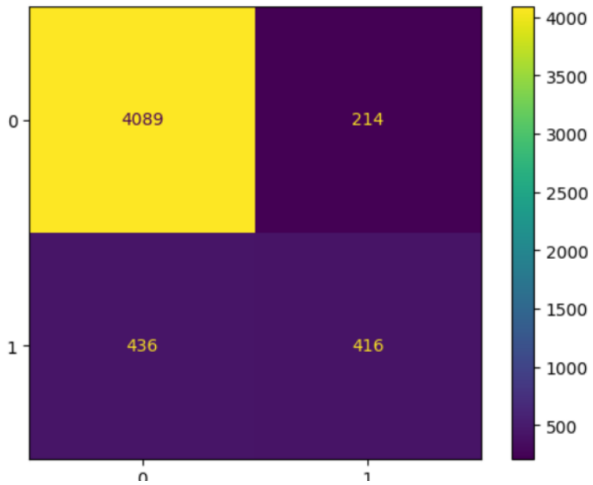
Training set accuracy *88.8%*:

Confusion matrix:



test set accuracy *87.8%*:

Confusion matrix:



References to Existing Works

AttentionSiteDTI: Interpretable graph-based model focusing on protein binding sites for predicting DTI. Link: https://academic.oup.com/bib/article/23/4/bbac272/6640006

DeepAffinity: Hybrid CNN and RNN architecture using attention for affinity prediction. Link: https://arxiv.org/abs/1806.07537