A. Modality Embedding

Biomedical Knowledge Graph

Node Attributes

Pretrained LMs

Attribute Embeddings

Gene/Protein
Nucleotide / Amino acid Sequence
Description

Drug
SMILES String / Amino Acid Sequence
Description

Disease
Definition
Description

ProteinBERT

DNABERT

BioBERT

MolFormer

B. Contrastive Learning

Fusion Module

GCN

Fusion Module

GCN

Fusion Module

GCN
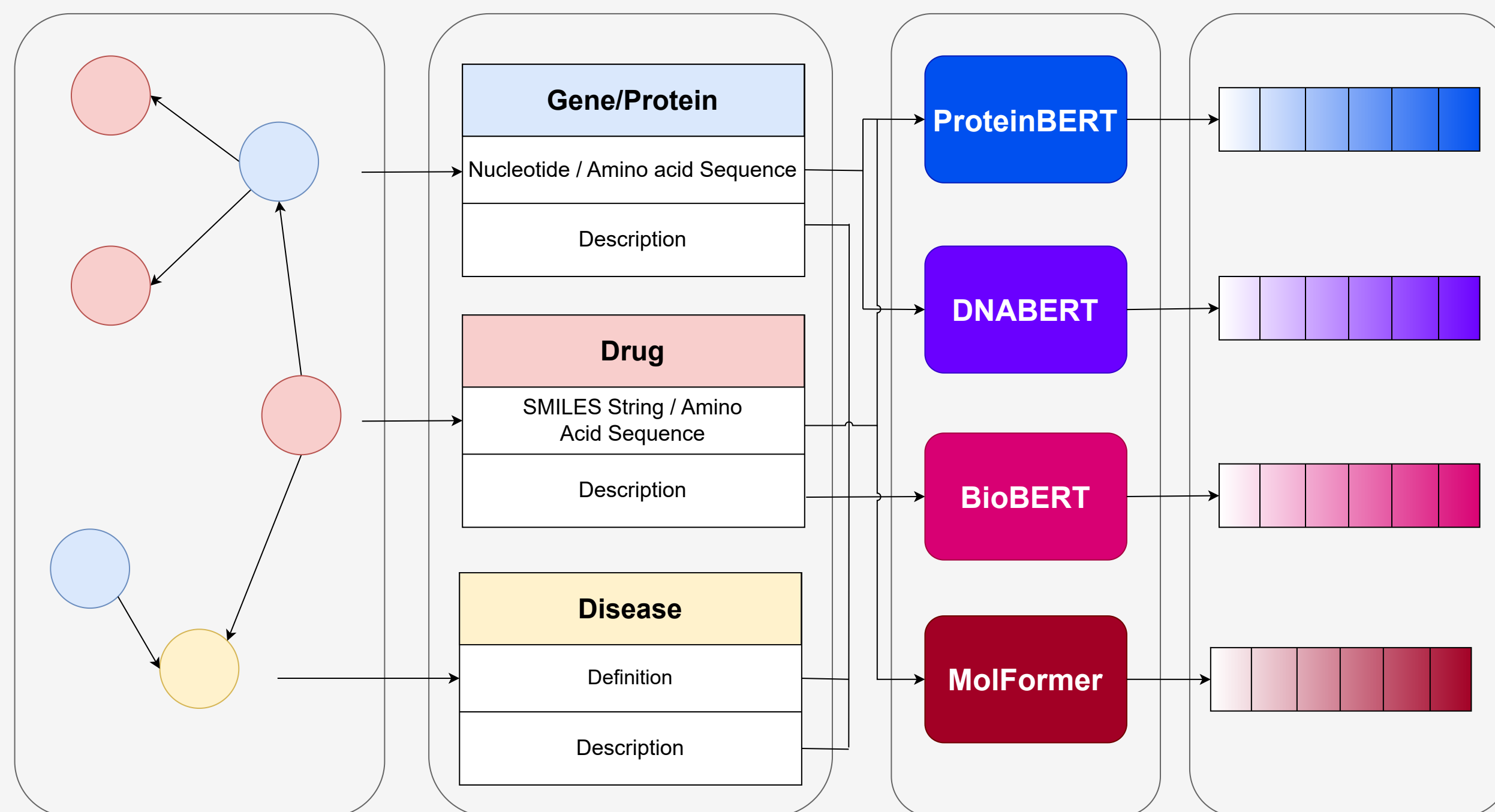
Contrastive Loss
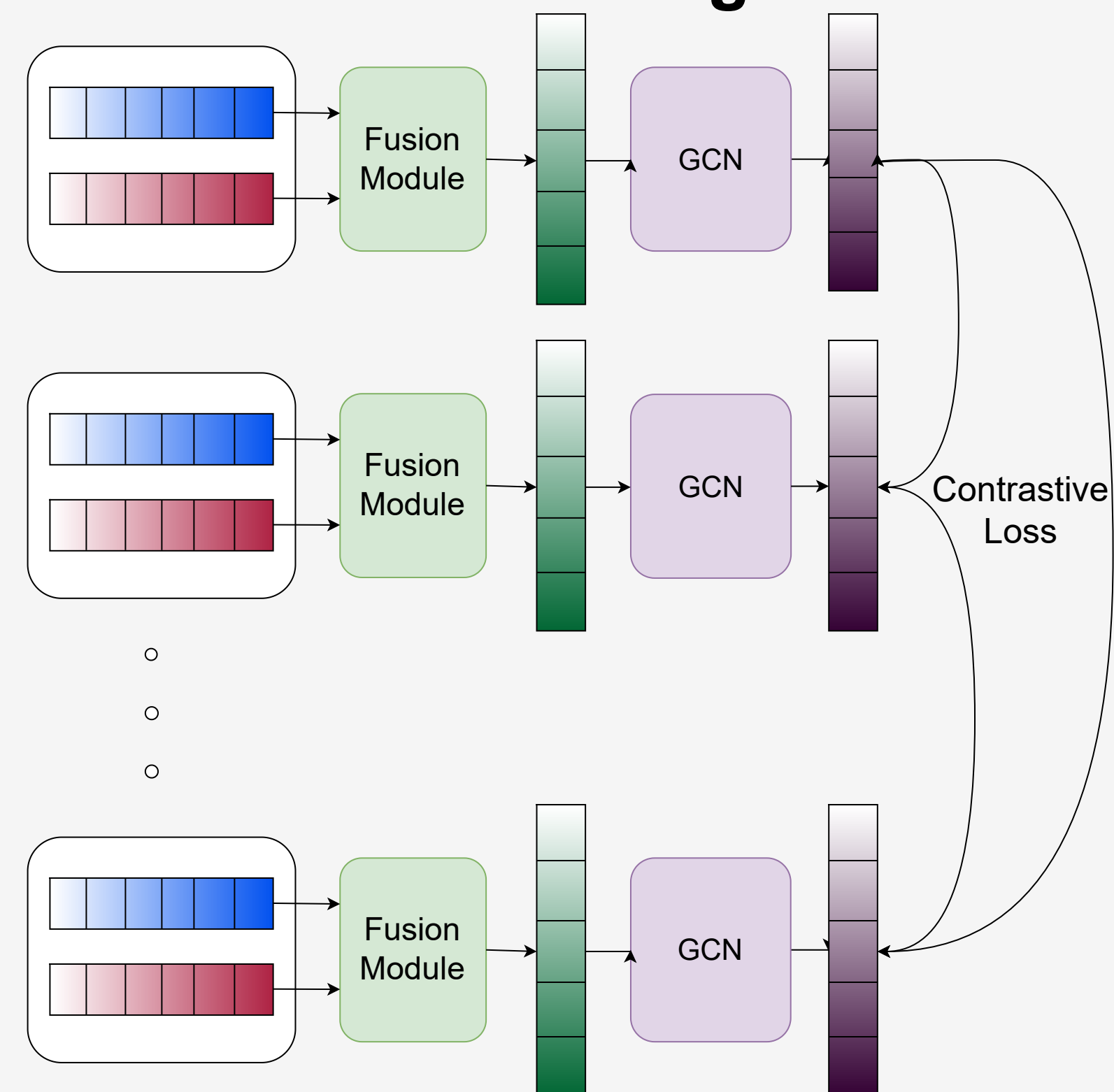
C. Link Prediction on KG Embedding

X_1 X_2 X_3 X_4 X_5 X_6 X_7 X_8 X_9
V_1 V_2 V_3 V_4 V_5 V_6 V_7 V_8 V_9

Input an incomplete graph with negative edges.

Hidden layer 1

Hidden layer 2

ReLU

Relational Graph Convolutional Network (RGCN) Encoder

Z_1
Z_2
...
Z_n

Distmult

Output a completed graph with predicted connections.