GLaMoER: Node-Edge Fusion via Gated Late Mixture of Experts with Residuals for Drug-Disease Association

Gusti Ayu Putri Saptawati, Dimmas Mulya, Hubbi Nashrullah Muhammad
Institut Teknologi Bandung, Indonesia

STEI-ITB

INTRODUCTION

Link prediction is a graph learning task that aims to predict missing or potential relationships between entities in a graph. In heterogeneous graphs, this task becomes important because multiple node and relation types can be modeled in one unified structure. Many link prediction methods rely primarily on node embeddings to score candidate links, which may overlook relational information specific to each pair. Therefore, this study propose GLaMoER (Gated Late Mixture of Experts with Residual), a model that fuses independently pretrained node and edge encoders. Gated late fusion combines the prediction scores from the two branches, and a mixture of experts refines the fused score by adding a learned residual correction. The experiments are conducted on the B-Dataset using Top-15 filtering and 5-fold cross-validation.

RESEARCH GAP

Most existing models focus on node-level representation, while link prediction requires explicit edge-level relational information.

RESULT

The standalone node and edge models showed complementary. TThe node-based model achieves the best sensitivity, with Recall = 0.720 and AUPR = 0.521. Meanwhile, the edge-based model is more selective, with Accuracy = 0.851, Precision = 0.376 and Specificity = 0.901.

GLaMoER with four experts (***K***=4) and a residual scale of 0.35 (***R***=0.35) produced the highest AUROC, Accuracy, F1-Score, Precision, and Specification compared to other combination of hyperparameter.

More Importantly, these settings win over five of seven metrics from Standalone Node and Standalone Edge model.

Model	AUROC	AUPR	Acc	F1	Prec	Rec	Spec
Standalone Node (LLM-DDA)	0.778	**0.521**	0.755	0.402	0.279	**0.720**	0.759
Standalone Edge (NIEE)	0.794	0.390	0.851	0.415	0.376	0.462	0.901
GLaMoER (K=4 , R=0.35)	**0.810**	0.433	**0.859**	**0.447**	**0.407**	0.495	**0.907**

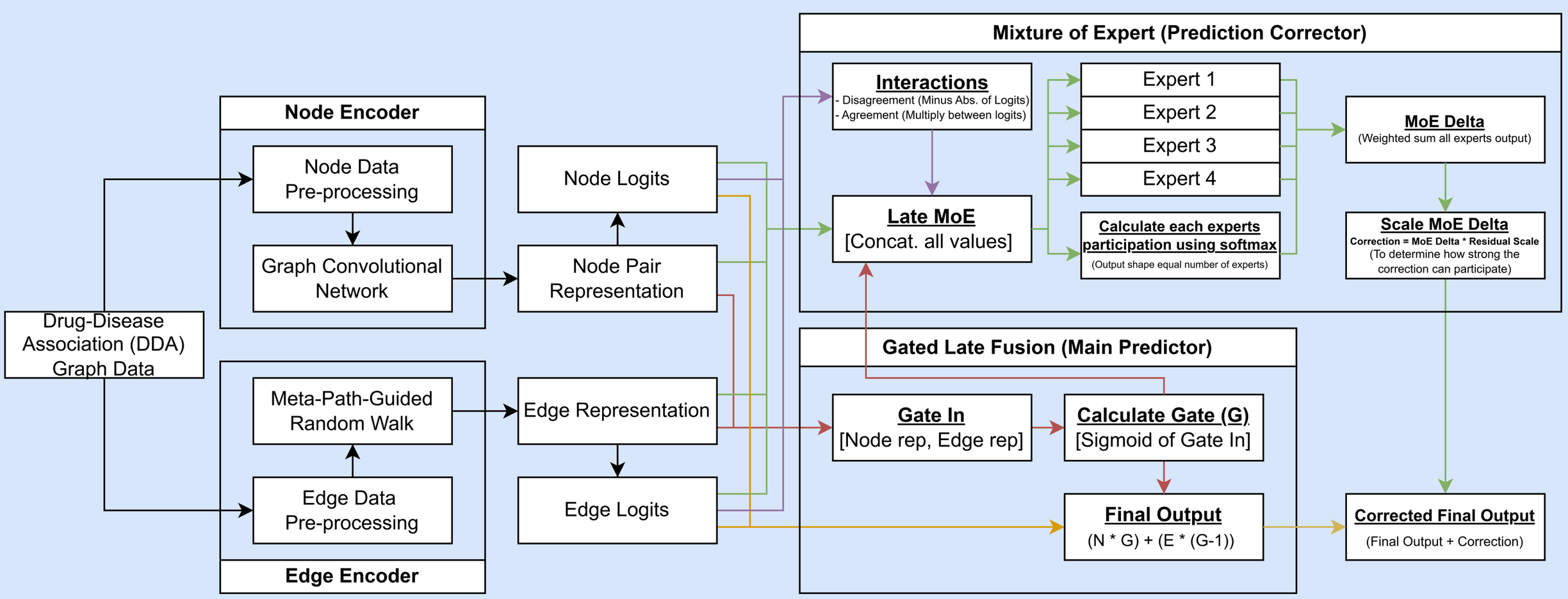
PROPOSED SOLUTION

The node encoder adapted from LLM-DDA and edge encoder from NIEE are pretrained independently within each fold. Their parameters are frozen during fusion training. For each drug–disease pair, a gate uses both representations to combine the node and edge logits. A mixture of experts then calibrates the fused logit by adding a weighted, pair-specific residual correction. ***K*** denotes the number of experts, and ***R*** controls the scale of the correction.

KEY CONTRIBUTION

This study proposed GLaMoER (Gated Late Mixture of Experts with Residual), a model with residual Mixture of Experts module for gated node-edge fusion and evaluates how the number of experts and the residual scale affect prediction performance.

GLaMoER: Gated Late Mixture of Experts with Residual



CONCLUSION:

Node and edge embeddings complement each other in link prediction. The standalone node model recovered more known associations, and the standalone edge model produced fewer false positives. GLaMoER achieved the highest AUROC and F1-score among the evaluated models with four experts and a residual scale of 0.35. Its gains over Gated Late Fusion suggest that refining the fused score with a residual mixture of experts can improve drug-disease association prediction.