A GENERATIVE CAUSAL NETWORK FRAMEWORK FOR MULTI-TASK LEARNING ON GRAPH-STRUCTURED DATA

Author’s: M. Premraj, K. Sujith, N. Vanjulavalli

Abstract

Graph-structured data are increasingly prevalent across domains such as social networks, bioinformatics, recommender systems, transportation systems, and financial fraud detection. While recent advances in Graph Neural Networks (GNNs) have significantly improved predictive performance, most existing models rely heavily on correlation-based learning and lack explicit modeling of underlying causal mechanisms. This limitation reduces robustness, interpretability, and generalization under distribution shifts. To address these challenges, this study proposes a Generative Causality-Driven Network (GCDN) for Graph Multi-Task Learning, a unified framework that integrates causal inference and generative modeling within a graph neural architecture.The proposed model assumes that observed graph structures and node features are generated from latent causal factors governed by a structural causal model (SCM). These latent variables capture key generative influences such as community strength, feature drift, and noise components, which jointly determine graph topology, node attributes, and task-specific targets. The architecture consists of four core components: (1) a Graph Convolutional Network (GCN)-based encoder to learn node and graph-level embeddings; (2) a causal inference module that estimates latent causal representations; (3) a generative decoder that reconstructs graph adjacency matrices and node features from inferred causal variables; and (4) multiple task-specific heads that simultaneously perform graph classification, graph regression, and node classification. By incorporating generative reconstruction losses alongside supervised objectives, the model enforces structural consistency between inferred causal factors and observed graph data. Furthermore, causal supervision is introduced to align learned representations with ground-truth latent variables in synthetic settings. The framework also supports intervention-based analysis using the do-operator, enabling counterfactual reasoning to assess how changes in causal factors influence downstream predictions. This capability enhances model interpretability and robustness under distributional shifts. Experimental evaluation on synthetically generated causal graphs demonstrates that the proposed GCDN achieves strong performance across multiple tasks while maintaining accurate recovery of latent causal variables. Compared to traditional correlation-driven GNN models, the causality-aware generative framework exhibits improved stability, better generalization under interventions, and enhanced explainability of learned representations.