Large Language Model-featured Academic Recommendation System
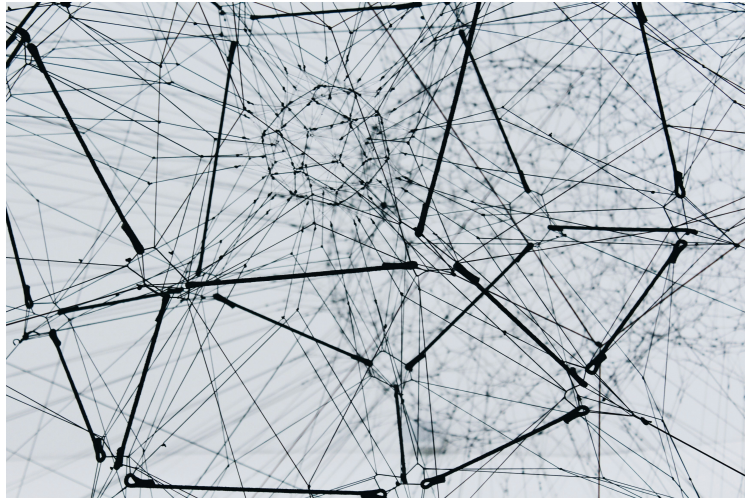
1 Objective

Considering the ever-increasing publication rate in academia, scholars in every discipline nowadays require more than ever before automated assistance to confront the threat of information overload. Current pipelines mainly utilize graph convolutional network(GCN) to capture the citation pattern, and hand-crafted methods such as skipgram or bag-of-words to generate textual features. Our scientific objective is to **improve the performance of link prediction** on Text-Attributed Citation Network with help of large language model and to generalize it to large scale graph with billions of nodes. In this pilot project, we aim to benchmark current possible approaches and seek useful insights for future work.

2 Approach

Specifically, we will benchmark three pipelines: **GCN without LLM (GCN), LLM as feature (LLM-Feat),** and **LLM as predictor (LLM-Pred)**. GCN includes all existing GCN methods, which can be roughly split into three types. LLM-Feat involves leveraging LLM to enhance node textual attributes and then generating predictions through Messaging Passing Neural Network (MPNN). The latter attempts to directly use LLMs as standalone predictive factors.

3 Contact

- Chen Shao (cc7738@kit.edu)

- Benjamin Schäfer (benjamin.schaefer@kit.edu)

Figure 1: Citation Network