Deep neural architectures for unsupervised graph representation learning

Popular science abstract

Motivation

When using various kinds of social media, such as Facebook or Twitter, we unknowingly become a part of a huge interaction network. More formally, such network is described using a so-called graph, which consists of nodes connected by edges. In the case of social networks, the nodes are the people using the media, whereas edges define different relations between the nodes, e.g., family relations, employment, viewed content or hobbies.
Based on such collected network data one could perform different tasks, including new friends recommendation or displaying personalized ads. Machine learning provides tools that allow to perform these tasks in an automated and relatively simple manner. Machine learning models require appropriate representations of the input data. Although some kinds of data do not need any additional preprocessing, in the case of graphs one must first build numerical descriptions of nodes, edges or subgraphs. The research area that explores these methods is called representation learning. It provides models and algorithms that can compute a vector representation (e.g., for each node) based on a given input graph. Most of the available algorithms is designed for graphs with a single type of edge, whereas many real-world datasets consist of multiple relation types (edge types).

Project objective

The main project objective is to close the gap of representation learning for multirelational graphs. The developed algorithms should provide an effective way to compute vector representations in multirelational graphs; and hence they should simplify model training in the typical graph processing downstream tasks.

Research description

Within this project's scope, two unsupervised representation learning methods for multirelational graphs will be developed. The first method will require building a loss function that will consider the existence of different relations in the graph and will exploit the available attributes (on the node, edge and whole graph-level). Additionally, there will be research conducted on how to utilize existing representation learning methods that were designed for simple graphs with one relation type. The second method will validate the research hypothesis that it is possible to infer the number of distinct relations types in a graph (based on the structure and available attributes). The last stage of the project will be an in-depth analysis of the proposed models and their performance in typical downstream graph representation learning tasks (e.g., node classification, link prediction or attribute regression).

Expected results

The main expected result of this project will be two articles presenting the described methods that are published on the best international conferences (e.g., ICLR, NeurIPS, ICML, ECML).