GEM: A Python package for graph embedding methods

Palash Goyal$^1$ and Emilio Ferrara$^1$

1 USC Information Sciences Institute

Summary

Many physical systems in the world involve interactions between different entities and can be represented as graphs. Understanding the structure and analyzing properties of graphs are hence paramount to developing insights into the physical systems. Graph embedding, which aims to represent a graph in a low dimensional vector space, takes a step in this direction. The embeddings can be used for various tasks on graphs such as visualization, clustering, classification and prediction.

GEM is a Python package which offers a general framework for graph embedding methods. It implements many state-of-the-art embedding techniques including Locally Linear Embedding (Roweis & Saul, 2000), Laplacian Eigenmaps (Belkin & Niyogi, 2003), Graph Factorization (Ahmed, Shervashidze, Narayanamurthy, Josifovski, & Smola, 2013), HOPE (Ou, Cui, Pei, Zhang, & Zhu, 2016), SDNE (Wang, Cui, & Zhu, 2016) and node2vec (Grover & Leskovec, 2016). It is formatted such that new methods can be easily added for comparison. Furthermore, the framework implements several functions to evaluate the quality of obtained embedding including graph reconstruction, link prediction, visualization and node classification. It supports many edge reconstruction metrics including cosine similarity, euclidean distance and decoder based. For node classification, it defaults to one-vs-rest logistic regression classifier and supports other classifiers. For faster execution, C++ backend is integrated using Boost for supported methods.

GEM was designed to be used by researchers studying graphs. It has already been used in a number of scientific publications to compare novel methods against the state-of-the-art and general evaluation (Salehi Rizi, Granitzer, & Ziegler, 2017, Lyu, Zhang, & Zhang (2017)). A paper showcasing the results using GEM on various real world datasets can be accessed (Goyal & Ferrara, 2018). The source code of GEM is made available at https://github.com/palash1992/GEM. Bug reports and feedback can be directed to the Github issues page (https://github.com/palash1992/GEM/issues).

References


