

DSS Statistics Seminar

March 18, 2022, 12:00

<https://uniroma1.zoom.us/j/86881977368?pwd=SWRFcVFjMDZTa0lXZk05TE1zNm5adz09>

Passcode: 432940

Unsupervised whole graph
embedding methods and
applications

Mario Rosario Guarracino

Dipartimento di Economia e Giurisprudenza, Università degli Studi di Cassino e del Lazio Meridionale, Consiglio Nazionale delle Ricerche - Italy

Networks represent a powerful model for problems in different scientific and technological fields, such as neuroscience, molecular biology, biomedicine, sociology, social network analysis, and political science. As the number of network applications increases, so does a need for novel data analysis techniques. In many applications, the analysis focuses on a single network to cluster or classify its nodes or predict pairs of nodes that will form a link. In this talk, we focus on problems where a network is a statistical unit, and the analysis regards whole networks rather than their parts.

Methods for learning features on networks focus mainly on the neighborhood of nodes and edges. We review some of the existing methodologies and introduce Netpro2vec, an embedding framework based on representations of graphs based on empirical probability distributions. The goal is to use basic node descriptions other than the degree, such as those induced by the Transition Matrix and Node Distance Distribution, to describe the local and global characteristics of the networks. The framework is evaluated on synthetic and real biomedical network datasets and compared to well-known competitors. Finally, open problems and future research directions are highlighted.



SAPIENZA
UNIVERSITÀ DI ROMA