

Deep Graphs - Represent and Analyze Heterogeneous Complex Systems across Scales

Dominik Traxl

Abstract

From a scientific point of view, "Big Data" provides a great opportunity. But it also poses novel methodological challenges, particularly since more and more information originates in unstructured form. To analyze data, methods developed in disciplines such as probability theory, multivariate statistics, non-linear dynamics and machine learning are employed. Regarding the representation of systems, network theory has proven to be a powerful instrument. Yet, even in its latest and most general form (i.e., multilayer networks), it is still lacking essential qualities to serve as a general data analysis framework. These include, most importantly, an explicit association of information with the nodes and edges of a network, and a conclusive representation of groups of nodes and their respective interrelations on different scales. To overcome these challenges, I developed a theoretical framework - Deep Graphs - which is based on network theory. This framework incorporates groups of objects (supernodes) and their respective interrelations (superedges) into a self-contained network representation. Furthermore, potentially unstructured and diverse information is explicitly associated with the different (super)nodes and (super)edges. For these reasons, my framework is capable of acting as a go-between, joining a unified and generalized network representation of systems with the statistical tools of traditional fields, as well as the methods developed in the rising field of machine learning. I will first introduce this framework along with its software implementation (written in Python). After pointing out the theoretical implications of Deep Graphs, I will demonstrate its utility by applying it to rainfall, wildfire and land use data.