Overview of Social Network-Based Inference

In this section, we'll discuss predictive analytics in social networks. Specifically, we'll discuss how to predict or infer the behavior of a node based on the behavior of the other nodes in the network. Various challenges arise when doing this. First, the data is no longer IID, or independent and identically distributed. The IID assumption is often made in traditional analytics when building predictive models using regression, decision trees, or other techniques. In a social-network context, the IID assumption is clearly violated because the behavior of one node might influence the behavior of other nodes. Hence, the correlational behavior between the nodes needs to be considered.

Furthermore, procedures for collective inference are needed because the nodes mutually influence one another. In other words, inferences about nodes can affect each other. Finally, there is no easy separation between the training set and the test set. We cannot simply cut the network in two parts because all nodes are interconnected.

The goal of social-network-based inference is to model network behavior that cascades from node to node much like an epidemic. Think about news, opinions, rumors, cascading failures in financial markets, viral marketing, and so on. To simplify the inference process and make it computationally feasible for large networks, the Markov assumption is often made. The Markov assumption states that the behavior of a node in the network depends only on the behavior of its direct neighbors.

Social Network Analytics

Close