CS768 - Learning with Graphs **Autumn 2020**

Lecture 1: August 7

Instructor: Abir De *Scribe: Aneesh Shetty (170040022)*

1.1 Machine Learning Models for Graphs

Discrimative Models on Graphs

The task of inferring some labelled data based on current knowledge of the Graph.

- **Link Prediction** : Based on current snapshot of graph $G = (V, E)$, predict which edges will appear from $V \times V - E$
- **Node Classification** : Labelling the nodes of a Graph based on Node and Edge Features.
  Node Features : partial labelling or probability distribution on labels of the nodes
  Edge Features : encode similarity, contact or other features
- **Embedding Design** : Given a graph as an input, and generate feature vectors for each node, which can then be used in some downstream task (Deep neural networks). But these require a lot of data. The model basically learns a compression technique. Given a graph, the feature of each node is generated by looking at the graph with the perspective of that node, and then compress that information into a dense vector ($D << |V|$). Without enough data, the compressed vector can be noisy. Moderate performance across many domains.

Both Link Prediction and Node Classification are Domain Dependent Problems (e.g. the solution working for social network may not in general hold for disease transmission networks, or collaboration network). Thus, there has been a Feature Engineering based approach to these problems.
To curb this, Embedding Design methods have been studied so as to apply them to a larger domain of applications
Even with simple handcrafted features, traditional models give as good results as embedding/deep learning models on moderate sized graphs.

Generative Models on Graphs

Given a set of small graphs (topology of smaller graphs) or some structure of Graphs generate new Graphs from scratch.

- **Barabasi Albert Model**: Generate graph in such a way so that degree distribution obeys a power law (Degree distribution often follows power law in Social Networks and Citation Graphs).
  If we want to fit network using Barabasi Albert Model, then we compute degree distribution of that graph and fit the Barabasi Power Law graph (parameter estimation).
  But still rest of properties of Barabasi graph may not match the real graph (Key problem)

- **Kronecker Graph**: Given a small graph, apply edge generation procedure to get larger graphs
- **Deep Generative Models**

1.2 Graphs are Everywhere

- Social Networks
- World Wide Web
- Information Networks
- Transportation on Networks
- Protein Interactions
- Internet of Things

Application of some methods in one domain may not be directly useful in the other.. Heuristics and Generative models for each kind of Domain might be different and we will see some of those.

1.3 Link Prediction Problem - Introduction

Given a Graph G_t, predict G_{t+1} (evolution of edges) Out of the Non-Edges of G_t which of them will be evolved to edges in G_{t+1}.

Goals

- Finding Facebook Connections
- Recommendation of movies in Netflix
  (Edge evolution/Link Prediction in a Bipartite Graph of Movies and Users)
- Predicting LinkedIn Connections

Challenges

- Traditional ML methods deal with IID variables (like SVM)
- But Graphs are defined by relations.. and so the independence assumption is not always useful
  (e.g. Direct SVM for 0/1 classification will not help since this assumes that all edges are from I.I.D random variables)

Methods

- **Structural Signal Based Method**: Does not require ML.
  Computation of some score function for each edge and each non-edge based on some topological structures and connectivity properties of the graph (e.g. Common Neighbors)
  These methods also give accurate predictions (specially on social methods). We will also look at the mathematical rationale of the ML free scoring methods
- **ML based methods** (especially Supervised): Computation of scores for the edges and non-edges but in such a way so as to minimize some sort of a likelihood function, and learn a predictive model $P_\theta(y_{uv}|G)$
  ML based innovative supervised learning methods other than just linear predictors that are run fast and give accurate results

1.4 Other Topics

Network Embeddings

- **Motivation**: Automatic Feature Designing (supervised linear predictor was a handcrafted feature)
- Compressing the entire network view to one information vector. It cannot be done in an ad-hoc manner, and planning the model that does the compression and learning has to be done carefully based on application.

Generative Models

- Given a set of small graphs (topology of smaller graphs) or some structure of Graphs generate new Graphs from scratch.
  - Barabasi Albert Model
  - Kronecker Graph
  - Deep Generative Models

Information Diffusion Models

- Interaction in social network like twitter
  - Who will retweet a persons message, will the message be retweeted
  - Publicizing an offer, to which nodes should it be advertised so that the diffusion of information is maximum
- It has some psycological aspects which will be touched upon a bit.