Argomenti Social networks

- Characteristics of **complex social networks**
    - **Giant component**
    - **Power law** (Heavy tail) distribution (to model the degree distribution)
    - **Globally sparse Locally dense**
    - **Clustering Coefficient**
    - Homophily
    - Monotone properties and properties that hold with high probability

- **Graph models**
    - Random Graph Model - **Erdos Renyi** (G(n,m) and G(n,p))
    - Small world Model - **Watts Strogatz**
    - Preferential attachment model - **Barabasi Albert** (also known as graph growth model

- **Information diffusion**
    - The **bass model** for epidemics
    - **Linear threshold** model e **Independent cascade** model
    - **Submodularity** in influence maximization
    - **Causation**/influence vs **correlation**
        - **shuffle test** e **Edge-reversal** test

- **Densest Subgraph**
    - **Density/Sparsity**
    - **Charikar's Greedy Approximation** Algorithm with proof!

- **Community Detection**
    - (Normalized) Laplacian Matrix
    - **Sparsest cut**
    - **Conductance**
    - **Theorem on the eigenvalues** of the normalized Laplacian matrix (Number

of zero-valued eigenvalues equals number of disconnected components)
    - **Cheeger's inequality** (Bound the conductance with the eigenvalues, because calculating conductance is NP-hard)
    - **Sweeping** algorithm: using second smallest eigenvector to compute a cut that **approximates** the conductance of the graph
    - **Spectral Clustering**: using the eigenvectors of the normalized laplacian matrix to get a partition of the graph with k connected components

- **Node embedding** (Encoding nodes of a graph as low dimensional vectors that summarize their graph position and the structure of their local graph neighborhood)
    - **Encoder and decoder** (The goal is to minimize reconstruction loss
        - Shallow Embedding
        - Pairwise decoders
        - Node features and local graph based embedding. Something that is used for GNNs.
    - **Random Walk embeddings** - Approach to place nodes in similar embeddings if they occur on short random walks on the graph
        - **Deep walk and node2vec** (often use cross entropy loss to minimize error)
            - Deep walk uses hierarchical softmax
            - Node2vec uses noise contrastive approach. Also differentiates the random walk basing itself on hyper parameters
        - **LINE (Large-scale Information Network Embeddings)**
            - Hybrid approach: two steps: direct node optimization and random walk