Summary of the 1st half

Lin Guosheng
School of Computer Science and Engineering
Nanyang Technological University

Outline

- 1. Data basics
- 2. Clustering – basic methods
- 3. Link analysis: PageRank
- 4. Graph convolutional network
- 5. Similarity Search
- 6. Clustering – advanced methods
- 7. Graph community detection

- How to prepare for the exam?
  - Review the lecture slides
    - Review the latest slides
    - You can skip the non-examinable sections
  - Review the tutorial slides
    - Tutorial materials are important!
  - Closed book
  - Bring calculators
  - Equations are **not** given in the exam paper

- 1 Data basics
  - Distance and similarity calculation
  - Data normalization
- 2 Clustering – basic methods
  - K-means, K-means++
  - Hierarchical Clustering

- 3. Link Analysis: PageRank
  - PageRank
    - Flow equations
    - Power iteration method
    - PageRank – Google Matrix
- 4. Graph neural network
  - Graph convolutional network
    - Forward pass
    - Training loss functions

- 5. Similarity Search
  - Local Sensitive Hashing – Random Projection
  - Product Quantization (PQ)
- 6 Clustering – advanced methods
  - DBSCAN
- 7. Graph community detection
  - Louvain Algorithm
    - Phase 1: Modularity optimization using node movements