Social and Technological Networks

Rik Sarkar

University of Edinburgh, 2019.
Course specifics

• Lectures
  – Tuesdays 12:10 – 13:00
    • 1 George square G8: Gaddum lecture theatre.
  – Fridays 12:10 – 13:00
    • Geography 2.13.

• Exam: 60%, Coursework project 40%.

• Exam:
  – To be held in April/May

• TA: Lauren Watson.
Today

• Why study networks?
• Relations to machine learning
  – Why networks are important in machine learning and vice versa
• Course page
  – Notes, exercises and course materials
• Course structure and coursework project
• Prerequisites
• Programming
Network or Graph

- A set of entities or nodes: \( V \)
- A set of edges: \( E \)
  - Each edge \( e = (a, b) \) for nodes \( a, b \) in \( V \)
  - An edge \( (a,b) \) represents existence of a relation or a *link* between \( a \) and \( b \)
Networks exist everywhere

• What are some different types of networks?
Example: Social networks

- Facebook, Linkedin, twitter..
- Nodes are people
- Edges are friendships
- The network determines society, communities, etc..
- How information flows in the society
- How innovation/influence spreads
- Who are the influential people
- Predict behaviour
- Make recommendations
World wide web

• Links/edges between web pages
• Determines availability of information
• Important pages have more links pointing to them
• Network analysis is the basis of search engines
Computer networks

• What can we say about the internet?
• How reliable are computer networks?
Electricity grid

- Network of many nodes, redistributing power
- Critical infrastructure
- Failure can disrupt ... everything
- Small local failures can spread
  - Load redistributes
  - Trigger a cascade of failures
- Network structure is critical

From Barabasi: Network Science
Road network and transportation

- Mobility patterns of people
  - Location data
- Suggest bus routes
- Suggest travel plans
- Traffic engineering
- Increasing importance
  - More vehicles
  - Self driving cars
Linguistic networks

- Networks of words
- Show similarities between languages
- Show differences between languages
- Document analysis
Business and management and marketing

• Business
  – What makes a restaurant successful?
  – Nearby restaurants? Community of customers?

• Marketing/management
  – Who are the influential people in spread of ideas/products?
Other networks

- Chemistry/biology
  - Interactions between chemicals
  - Interactions between species
  - Ecological networks
  - Networks of neurons, blood circulation
- Finance/economies
  - Dependencies between institutions
  - Resilience and fragility
- Neural (Brain) networks
Network analysis in data science

• Data getting more complex
• Many types of data are not points in $\mathbb{R}^d$ space
  – Data carry relations – networks
  – Simple classification inadequate
  – Network knowledge can make ML more accurate, efficient
  – E.g. data from social network or social media, www, IoT and sensor networks, road networks
Machine learning

• Finding groups of points

• Separating groups of points
Machine learning

- Finding groups of points
- Separating groups of points
Challenge: ML on networks

• How do we do machine learning when the data is not in space, but in a network?
  – What does clustering mean?
  – What does separation mean?
Challenge: ML on networks

- Network data shows up everywhere
- We need to generalise ML to work on networks for more advanced operations
- The maths that works in Euclidean ($\mathbb{R}^D$) space has to be modified to work on graphs.
Networks for ML

• Another perspective:
  – Networks are useful for ML

• Example
  – Clustering with DBSCAN
    • Connect points that are close to each other to make graph
    • Take connected components to get clusters
  – Easily finds oddly shaped clusters

• Networks are good for determining the \textit{shape} of data
Networks for ML

- Used in
  - Clustering
  - Robotics
  - Motion planning
  - ....
Topics

• Networks, ML and algorithms
  – Community detection (clustering)
  – Predicting unknown values at nodes (classification)
  – Kernel methods
    • Graph kernels
  – Influence maximisation
    • Finding representative items and sampling

• Properties of common networks and models
  – Power law networks
  – Small world graphs (six degrees of separation etc)
  – Web graphs
  – Epidemics and cascades

• Theory, maths, statistics
  – Properties of random graphs and other common types of graphs
  – Metric spaces
  – Expansion, growth etc

• This is an advanced course to help research and innovation. We will try to balance between covering a range of interesting topics and studying them in depth
Web page

- Web page
  - http://www.inf.ed.ac.uk/teaching/courses/stn/
- Lookout for announcements on the web page
- Reading materials, slides, exercise sets will be uploaded to the web page.
Note and exercises

• Some material will be covered in lectures, other materials will be given as notes and exercises

• Please follow along as these are uploaded
  – Solutions to some exercises to be uploaded 1-2 week afterward

• Suggestion: Create your own study groups of 3 – 5 people and discuss
  – Try to write the solutions, proofs as cleanly and logically as possible
Lectures

• Please attend the lectures

• Bring notebook and pen
Coursework Project

• You will be given option of 10-12 projects
  – Pick one to do
• Topics expected in areas of:
  – Machine learning and optimisation
  – Algorithms and data structures
  – Data mining
  – Recommendation systems
  – Social networks
  – Linguistic networks and analysis of stories
  – Road networks or maps
  – Self driving cars (possibly)

  – Find your own topic!
Coursework Project

• You will be given a general topic area
• Your job is to:
  – Understand the domain and identify a question to answer. Determine its motivations.
  – Formulate the problem precisely, mathematically.
  – Find good solutions
    • Show that your solution works well
    • Can be theoretical or experimental (or both)
    • Either way, you need to be rigorous – be able say exactly where it works or does not work, and why
  – Write a good report that explains all of the above nicely
Project

- 1 project. 40% of marks
- Given: Around Oct 10 to 15.
- Due: Around Nov 15.
- **Objective:** Try something new in network science.
- Submit code and ≈3 page report
- The usual project consists of motivation, problem formulation, some mathematical/algorithmic ideas and verification by experiments

- We assume that you can program and run common algorithm and ML libraries
- Marked on
  - Rigor of work
  - Originality in problem and solution
  - Clarity of presentation
Projects

• Open ended projects are common in real world
• People that can do original work are highly valued in industry
• Your BSc/MSc projects are open ended
  – You are given a topic. You have to define exactly what to do and how
• A course project can help your BSc/MSc project
  – Network science, graph theory, are relevant to most CS areas
  – It is an opportunity to learn more about the area
Programming and python

• We will occasionally use python with jupyter notebooks in class
• Setup instructions on web page
• Sample notebook with lecture slides. Try it out!
Theory Exam

• Standard exam, 60% of marks
• Explain phenomena, devise mechanisms, prove properties...
• Last year’s paper online..
Pre-requisites

- See Topic 0: Background at
- Probability, distributions, set theory
- Basic graph theory and algorithms
  - Graphs, trees, DFS, BFS, minimum spanning trees, sorting etc
- Asymptotic notations
- Linear algebra
- Read up on all these materials and notations
- Do exercise 0

- Make sure you know this material
- And can do exercises 0 without help, and can explain your answers

- From next class, I will assume that you know this material.
Course learning expectations

• Plan and execute original projects
• Use programming for data driven analysis
• Use theoretical analysis to understand ideas/models rigorously
• Present analysis and ideas
  — Precisely
  — Unambiguously
  — Clearly

• Have fun playing with new ideas!