IST 402 Week 12 Notes

Zach Ripka, Jimmy Steff, Brennan Cornwall

11/12

AI for social good

a. Despite issues with ML, AI has many positive impacts
b. How can AI machine learning solve societal problems?

How can we use AI to solve low resource communities problem?

a. No access to public health facilities
b. Healthy food
c. Access to rehab facilities
d. Access to good governance (farming communities)

Many of the problems of low resource communities can be solved by increasing social networks (*not necessarily online*)

Current Problems

Typically NGO's who help these communities don't have enough money, time and volunteers

How can AI be used to raise awareness of HIV prevention to homeless youth?

- Homeless use 10x more likely to get affected by HIV

Traditional method:

1. NGO’s conduct “interventions” peer leader is given information and taught how to spread this information to their peers.
2. Urge them to adopt safer behaviours
3. Encourage to spread message
4. Diffusion of information
5. Resource constraints, cannot intervened on every homeless youth themselves.
6. Rely on word of mouth
7. GOAL: maximize number of youth who is informed about HIV

1. Friend network, which people should we target to maximize spread of information (similar to viral marketing)

Standard Influence maximization

a. Inputs:
    i. Social network (G)
    ii. Influence model (I): description of how influence spreads between people
    iii. Number of Nodes to choose (K): depends on number of people in homeless shelter
b. Output
    i. Optimal set of K nodes to maximize expected influence spread
c. INFLUENCE maximization algorithms are not great with real problems

Critical challenges in the real world: uncertainties

a. Uncertainty in social network structure
    i. Homeless don’t use social media (hard to gather data)
    ii. Uncertain about existence of
b. Uncertainty about state of influence of nodes
    i. Unable to sample every node
    ii. Unknown which nodes are influenced and which are not
c. Multi-stage selection
    i. Does not work for influence maximization

Given a social network G = (V,E)

a. Picking M subsets of nodes (M interventions organized by shelter) 3
b. Size of each subset is K (Max capacity of shelter) 4 leaders

2 algorithms to solve this problem, HEALER and DOSIM

a. 3 studies with 173 homeless youth in LA
b. 2 shelters (my friends place and Safe place for youth)
c. Raise awareness about HIV
d. Goal: verify need for using AI algorithms, verify usability

HEALER/DOSIM/ Degree Centrality - select peer leaders

a. Network application: interacts with homeless youth
b. Algorithm: determines which people should be recommended for the intervention
c. Shelter official collects information and refines algorithm (are people really friends or not)
d. Looks at social network data and selects a set of nodes for first intervention
e. What percentage of non peer leaders were informed about HIV after the interventions 70% healer and dosim
f. What percentage of informed, non peer leader, untested people made a proactive move afterwards and got an HIV test
g. Improvement of 150% over degree centrality

Degree Centrality (selects most popular people)

a. Creates too many redundant edges
b. Influencing people who are already influenced
c. Cross community edges are poor
d. Ignores cross community connections (only need to influence 1 person in a small group)

What is wrong with the study

a. 3 different populations
b. Sample size too small
c. Randomness
d. Only ran the algorithm once
e. Can’t run twice
f. Can't determine statistical significance
g. Different messages spread differently

11/14/19

1. Network application (connects with youth) Algorithm (Feedback)
2. # of non leaders informed: Healer and Dosim 70% DC 27%

Maximizing the spread of influence through social network

a. How to mitigate/maximize spread of influence

Influence/Scenario

1. Influence: can cover anything (awareness/product/propaganda)
2. Scenario Given: a limited budget B for advertising, estimate **influence between individuals?**
   a. Goal: trigger a large cascade of influence
   b. Question : which set of individuals should B target
   c. **What you need:**
      i. Form models of influence in social networks
      ii. Obtain data about particular network (estimate inter personal influence)
      iii. Devise algorithm to maximize spread of influence)

Viral marketing: goal - ensure further adoptions of a product

S- curve of adoption: the spread of innovations curve

a. Y: number of adopters
b. X: time S - top part: sold too much

Applications besides product marketing

a. Spread innovation
b. Detect stories in blogs

Models of influence

a. Two basic classes of diffusion models: threshold and cascade
b. Operational View
    i. A social network is represented as a directed graph and each node is a person/customer
    ii. Nodes start either active or inactive (influenced or not)
    iii. An active node may trigger activation of neighboring nodes
    iv. Monotonicity assumption: active nodes never deactivate
    v. Directed vs undirected graph (directional edges)
    vi. Bidirectional normal friendship

Linear model threshold

a. A node v has random threshold (0.7, will not become influenced until 70% of people in social network become influenced)
b. Every relationship as different influential threshold number on a person (person around them .3 .5 .2 etc..)

Independent Cascade

a. Probability that if you get influenced the other person will get influenced
b. Person you rarely see, low percentage -- good friend, high
c. Strength of friendships

How to solve influence maximization problem

a. Set of K-nodes (initial adopters)
b. Greedy algorithm
    i. Repeat procedure for K timesteps
    ii. For each timestep: add a single node to the result set
    iii. Results - 4 nodes
c. Also works with any algorithm that has diminishing returns property(return on investment) math name - submodularity
d. Influence maximization problem with diminishing returns property means always use the greedy algorithm

Real world algorithms: just give ranges of nodes from each other

a. Rank people based off of the discussion of their friends, how good of friends is the other person
b. Use ranges (be robust)
c. Assume there is

3 ways games can be solved

a. Nash Equilibrium - what we are using
b. Dominance
c. Something else