SAND2022-7747C

Sandia National Laboratories

Exceptional service in the national interest

Simulation Testbed

PRESENTED BY

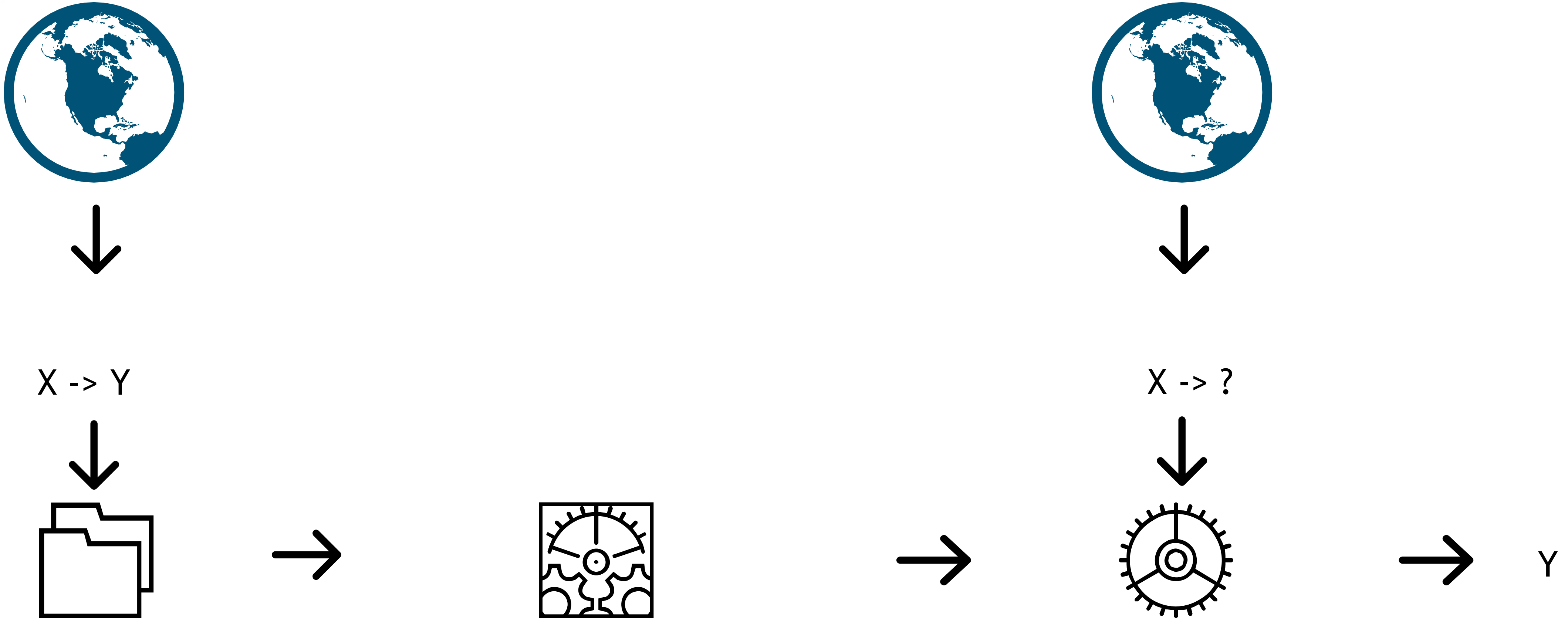
Presented by: Asmeret Naugle, 1462

Team: Kiran Lakkaraju (PI), Matthew Sweitzer, Steven Winyadinata, Griffin Lehrer

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. SAND20XX-XXXX P

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

STANDARD MACHINE LEARNING PROCESS



Relevant training data about the problem is sampled from the real world.

- Disinformation tweets, retweets.

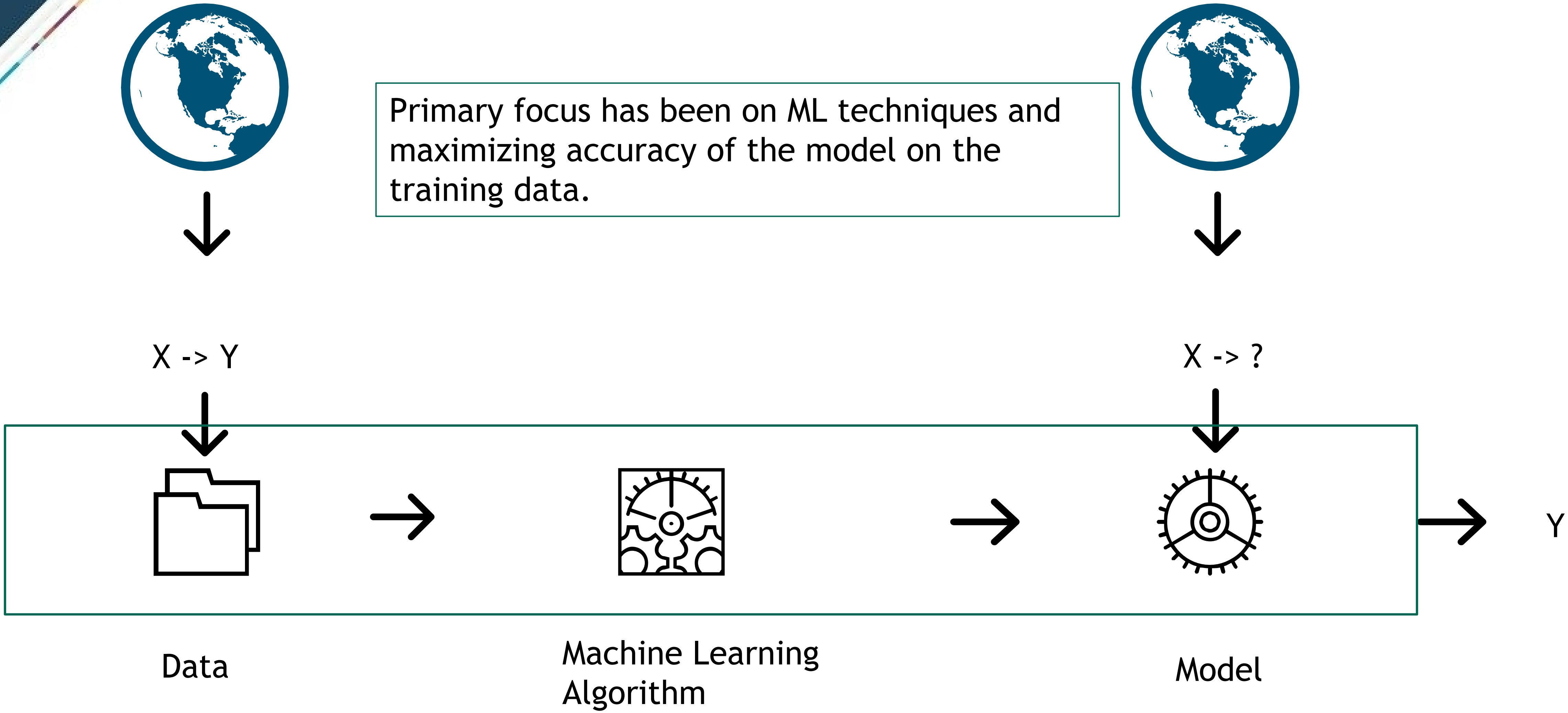
Patterns in data are identified using machine learning/artificial intelligence techniques.

- Deep Learning
- Bayesian networks
- Causal reasoning

A model is used to help predict real world patterns.

- If a tweet has pictures (X) -> it will go viral (Y)

STANDARD MACHINE LEARNING PROCESS



HOW DO WE HANDLE A CHANGING WORLD?

Problem: How do we handle a changing world?

- Model may have been trained on “stale” data.
- Model may make incorrect predictions.

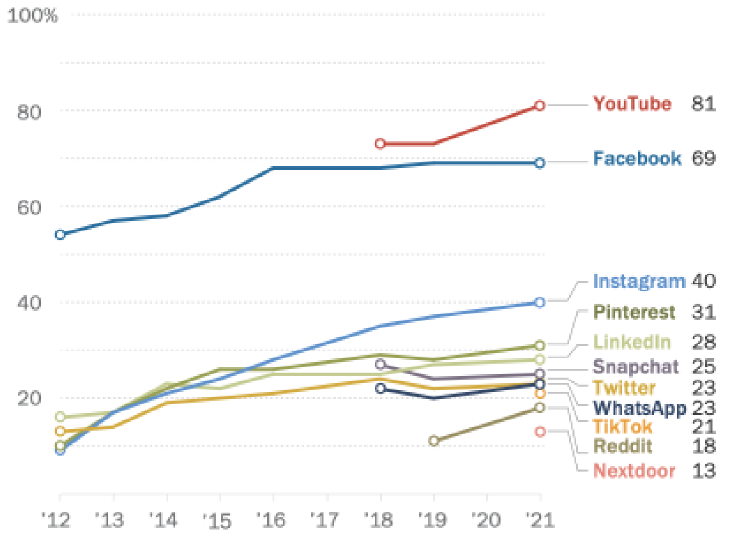
- Problem: Data may have sample bias and class imbalances.

SMALL CHANGES IN THE WORLD CAN CAUSE UNKNOWN EFFECTS

- Social media companies change user elements all the time.
  - Twitter changed the length of tweets.
  - Facebook included “frowny face” emojis
- Social media companies change invisible elements as well.
  - Facebook and Twitter friend/follower recommendation algorithms.
  - News/message recommendation.
  - Search results (based partly on auction).
- The world changes
  - Demographics of usage shift.
  - New technologies emerge (snapchat/tiktok vs. facebook)

Growing share of Americans say they use YouTube; Facebook remains one of the most widely used online platforms among U.S. adults

% of U.S. adults who say they ever use ...

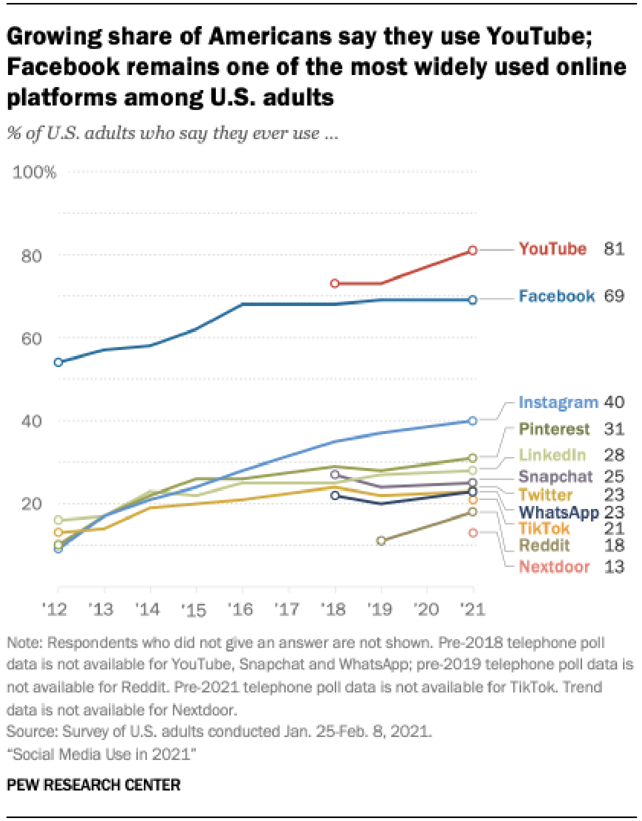


Note: Respondents who did not give an answer are not shown. Pre-2018 telephone poll data is not available for YouTube, Snapchat and WhatsApp; pre-2019 telephone poll data is not available for Reddit. Pre-2021 telephone poll data is not available for TikTok. Trend data is not available for Nextdoor.
Source: Survey of U.S. adults conducted Jan. 25-Feb. 8, 2021.
“Social Media Use in 2021”

PEW RESEARCH CENTER

HOW WELL DO MODELS ACCOUNT FOR THESE DIFFERENCES?

- Social media companies change user elements all the time.
  - Twitter changed the length of tweets.
  - Facebook included “frowny face” emojis
- Social media companies change invisible elements as well.
  - Facebook and Twitter friend/follower recommendation algorithms.
  - News/message recommendation.
  - Search results (based partly on auction).
- The world changes
  - Demographics of usage shift.
  - New technologies emerge (snapchat/tiktok vs. facebook)
- Changes have UNCERTAIN impact.
  - Example: surprising result from famous facebook experiment.



How do we assess models for resilience to changes in the world?

SYNTHETIC DATA IS GROWING IN USE

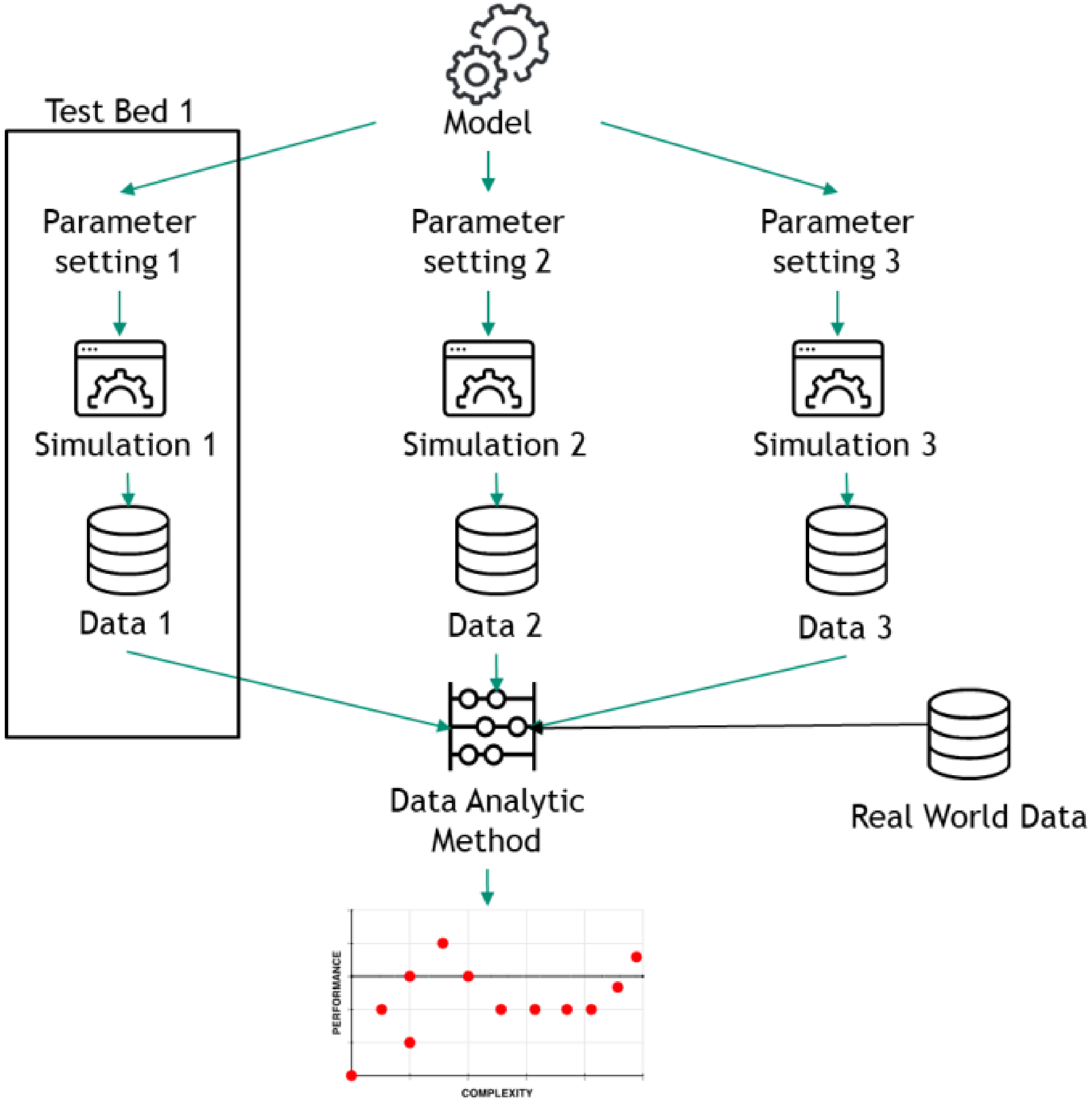
- Synthetic data is being used more in machine learning.
- Training Autonomous cars, facial recognition.
- Most of the uses are around machine vision applications.
  - Generating new scenes etc.
- Synthetic data for social media focuses on graph topology.
- Most methods focus on historical data and matching high level metrics.
- Current methods do not account for novel changes in the world.

WE WILL GENERATE SYNTHETIC DATA FROM SOCIAL SIMULATIONS

Generate synthetic social data with known ground truth by using social simulation methods.

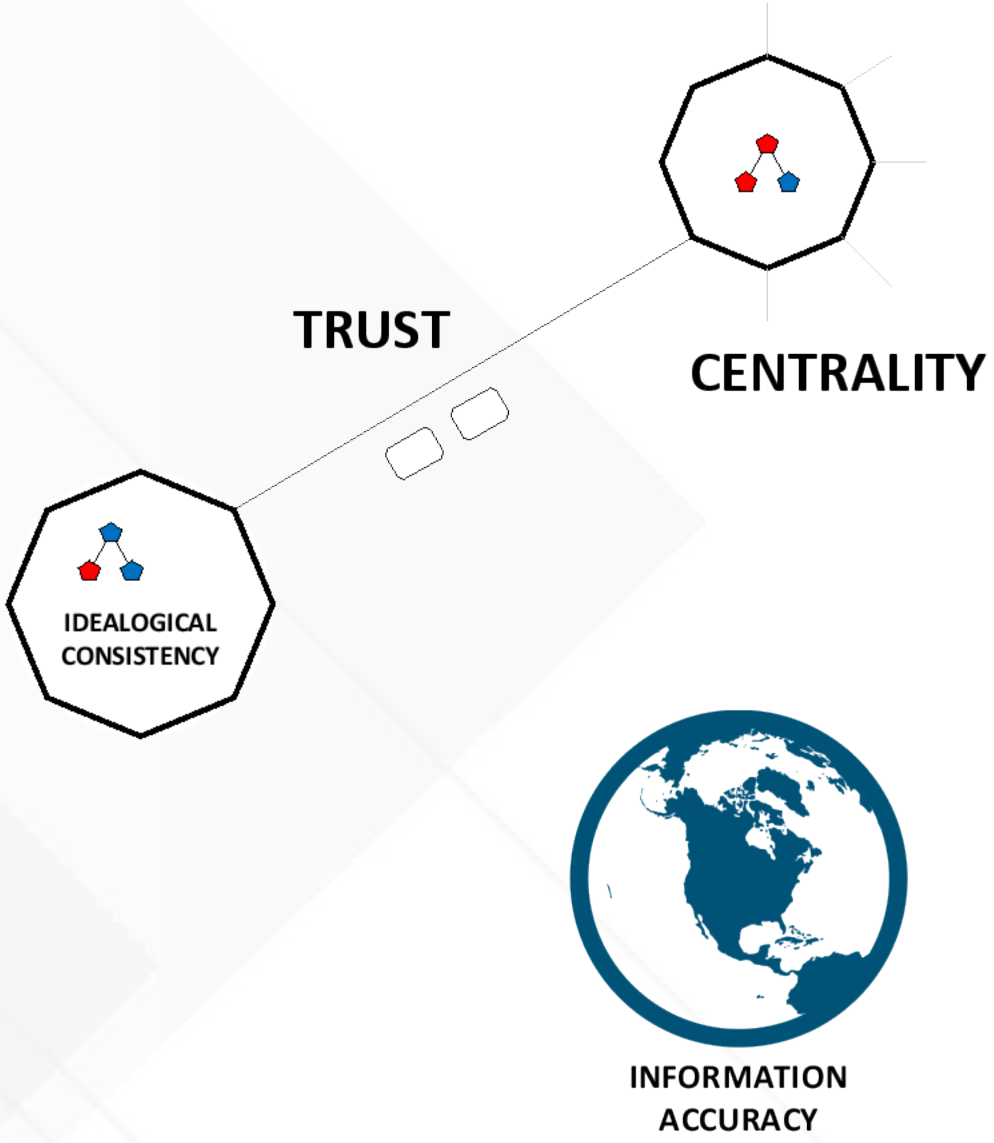
- Social simulation methods using agent-based modeling to capture underlying cognitive and social features.
- Incorporate important elements drawing from social science theory.
- Known ground truth – we know exactly how complex the underlying model is.
- Key research questions:
  - *RQ1:* How does complexity impact the effectiveness of data analytic methods?
  - *RQ2:* What simulation test bed characteristics correspond to real-world performance?
  - *RQ3:* How well do common data analytic methods perform under simulated concept drift?

WE ARE DEVELOPING A PIPELINE



WE HAVE A MODEL OF DISINFORMATION FLOW

- Agent based model of disinformation flow.
- Key components:
  - Agents with internal ideologies.
  - Message passing between agents
    - Similar to Twitter or email.
  - Agents can reshare incoming messages.



RQ3: HOW DO ALGORITHMS PERFORM UNDER SIMULATED DATA DRIFT?

- Outline of method:
  - Vary parameters of simulation.
    - Parameters that were varied:
      - P1..P2
  - Train a machine learning model using an algorithm
    - M1: Linear regression
    - M2: Decision trees
  - Train the algorithm using data from PX, then test on PY
    - This simulates data drift – a change in the underlying parameters settings.
    - Predictions:
      - Some parameter regimes will be more difficult than others.
      - M2 will perform better under a wider range of parameter settings than M1.

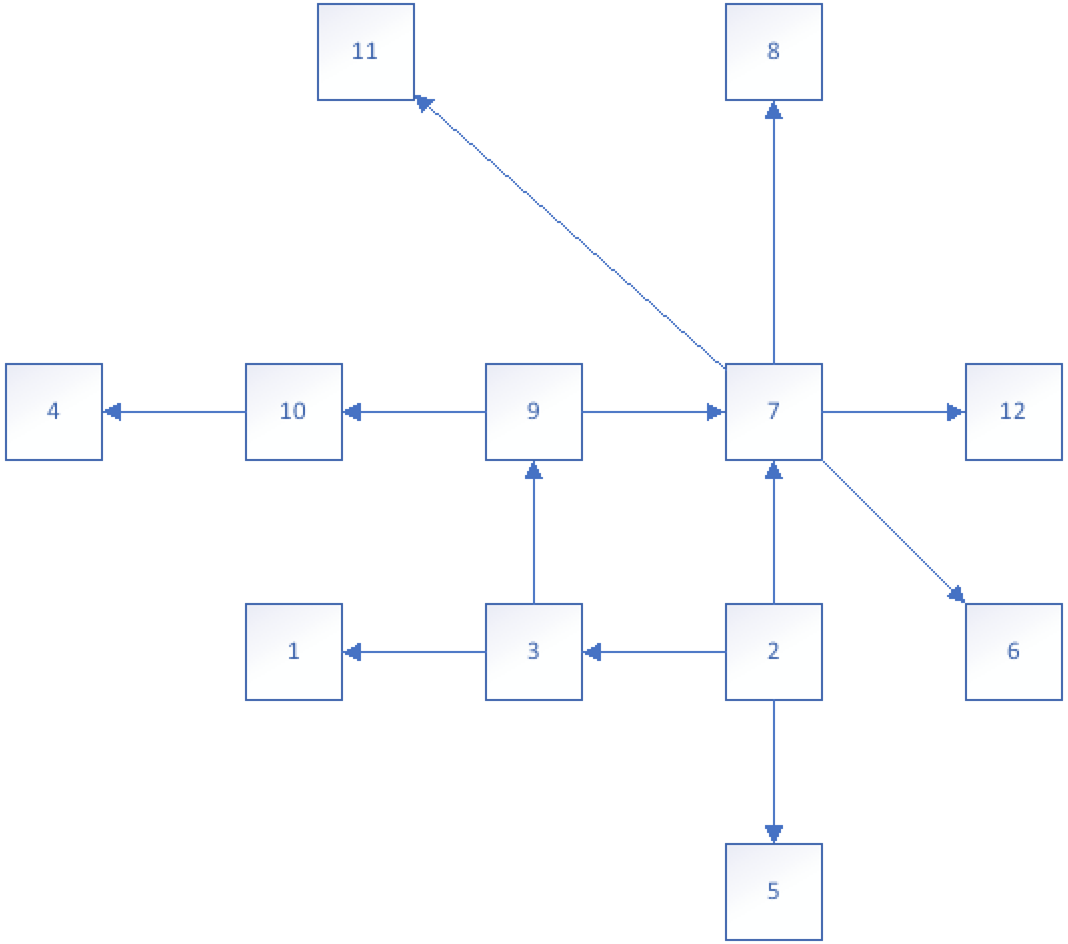
NOTATION TO HELP CLARIFY

- S_i = Data from a social simulation with parameter setting i. S_i^{train} is the training set, and S_i^{test} and S_i^{val} are the test and validation sets.
- $M_i = train(S_i^{train})$ Mi is a model trained on data generated from social simulation S_i.
- $E = Predict\left(M_i, S_j^{val}\right)$: The performance of using M_i to predict based on data generated from social simulation S_j.
  - i.e., how well does a model perform when its used on data from a different "world".
  - Note that the dependent variable may be the same, but fundamental elements of the simulation may differ between S_i and S_j.
  - For instance, one could be a situation where the social network is scale free, and Sj a situation where the social network is small world.
  - Error will be defined per problem domain/model type.

POPULARITY PREDICTION IS OUR PROBLEM DOMAIN

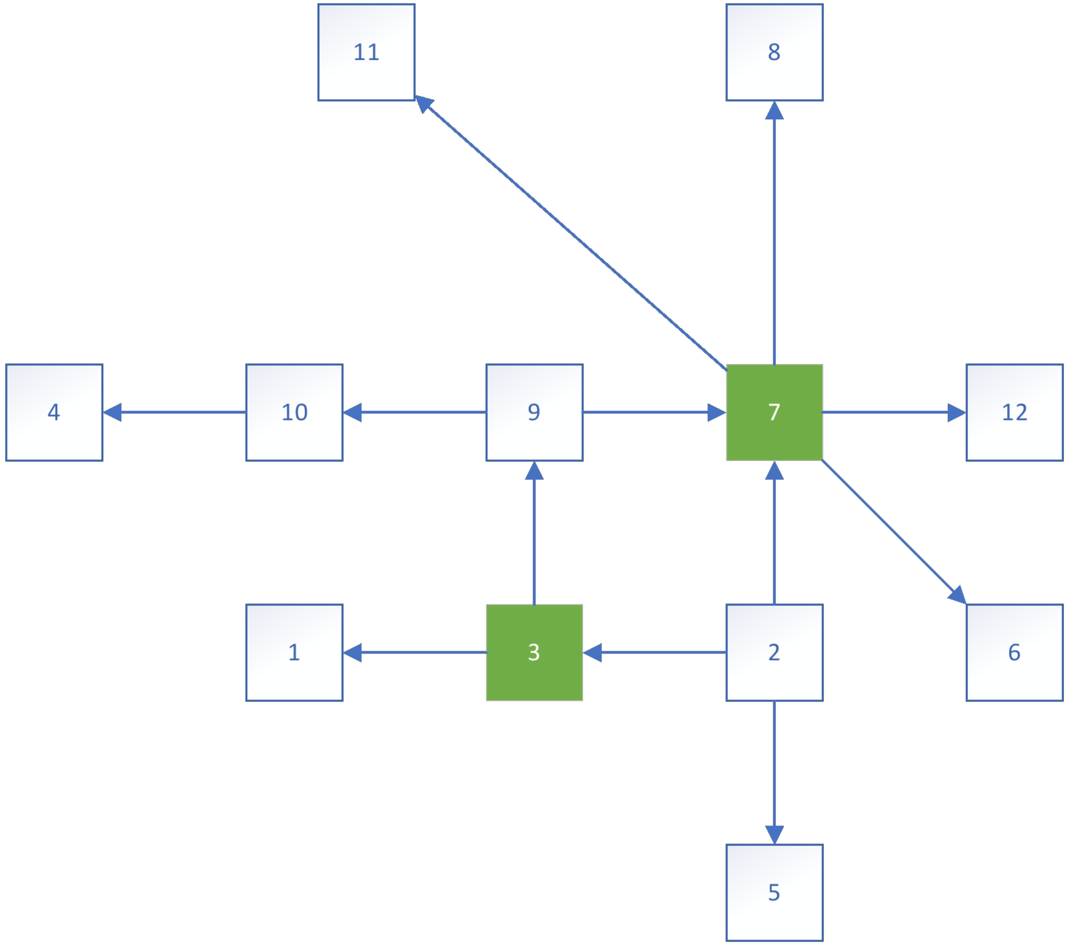
- Problem: Predict the popularity of information using only the first few timesteps of data.
  - Example: A disinformation tweet emerges 2 days ago. We need to predict how popular this tweet will be in a week so we can decide if mitigation measures are needed.
- This is called popularity prediction.
- What we have:
  - Information about the progress of the tweet through the social network.
    - For example: B retweeted post i from A at time 54.
- What we want to predict: How many people retweeted a post x number of timesteps after the

SETUP: SOCIAL NETWORK



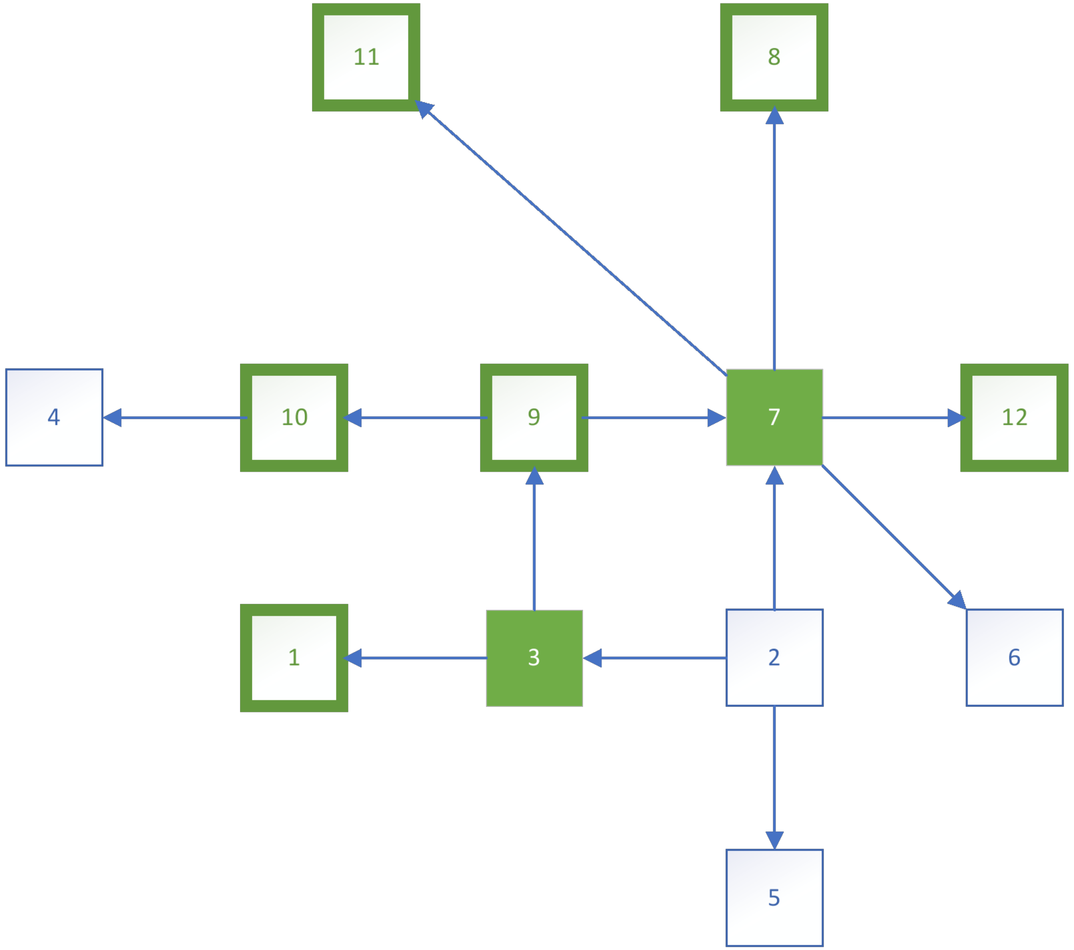
SETUP: CASCADE

- Cascade starts with initial distributors – “authors”.

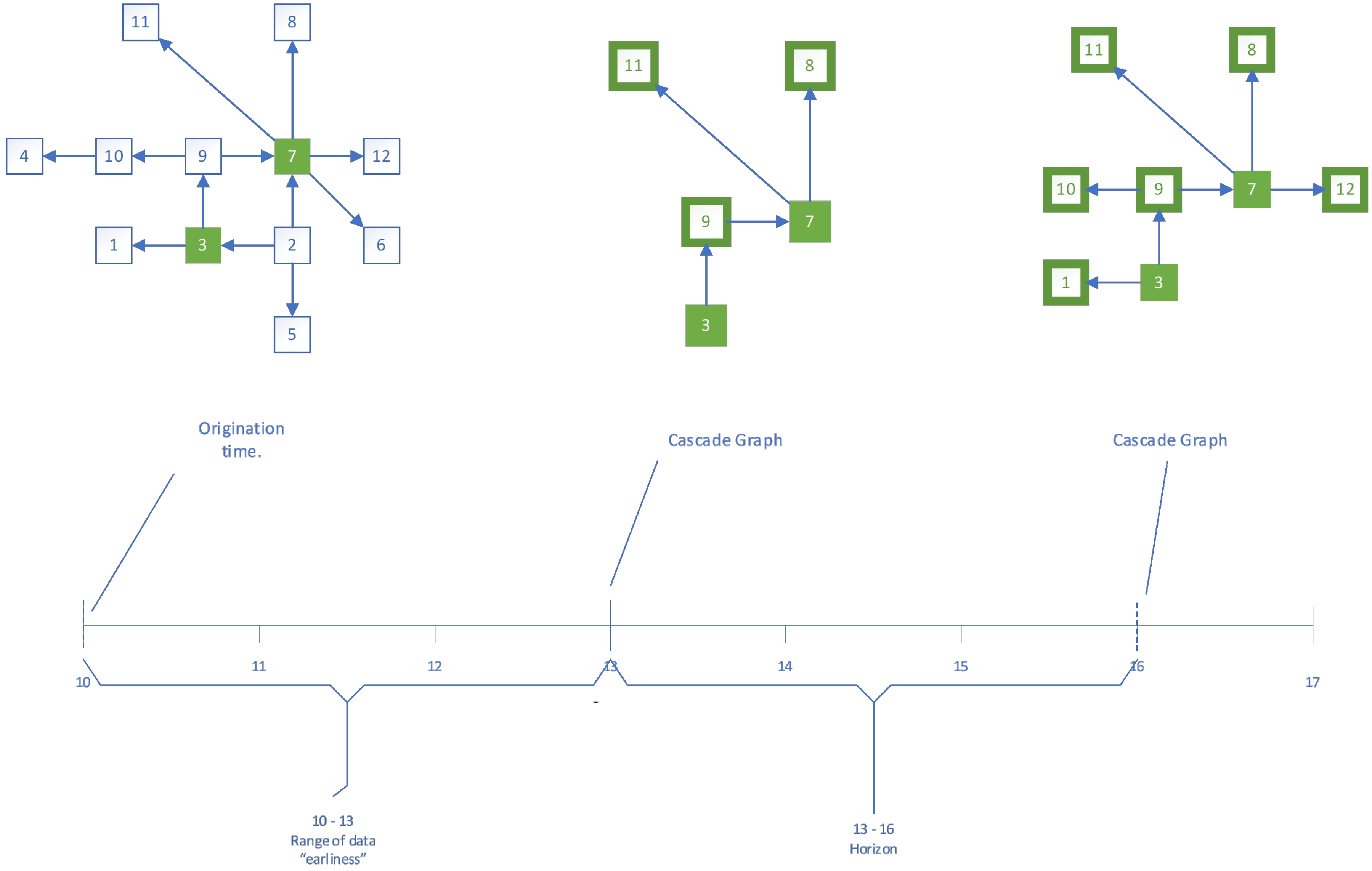


SETUP: CASCADE

- Cascade ends after a certain time period.
- Are there nodes in the cascade that are NOT linked to the authors?
  - There is no path from an author to this person.



CASCADE PREDICTION



CASCADE PREDICTION PROBLEM

- Given V_c^t as the number of nodes that adopted cascade c by time t.
- Goal is to predict:

$$\Delta s_c = |V_c^{t+\Delta t}| - |V_c^t|$$

- i.e., predict the additional number of adoptions that occur within timesteps.

SIMPLE INFORMATION DIFFUSION MODEL

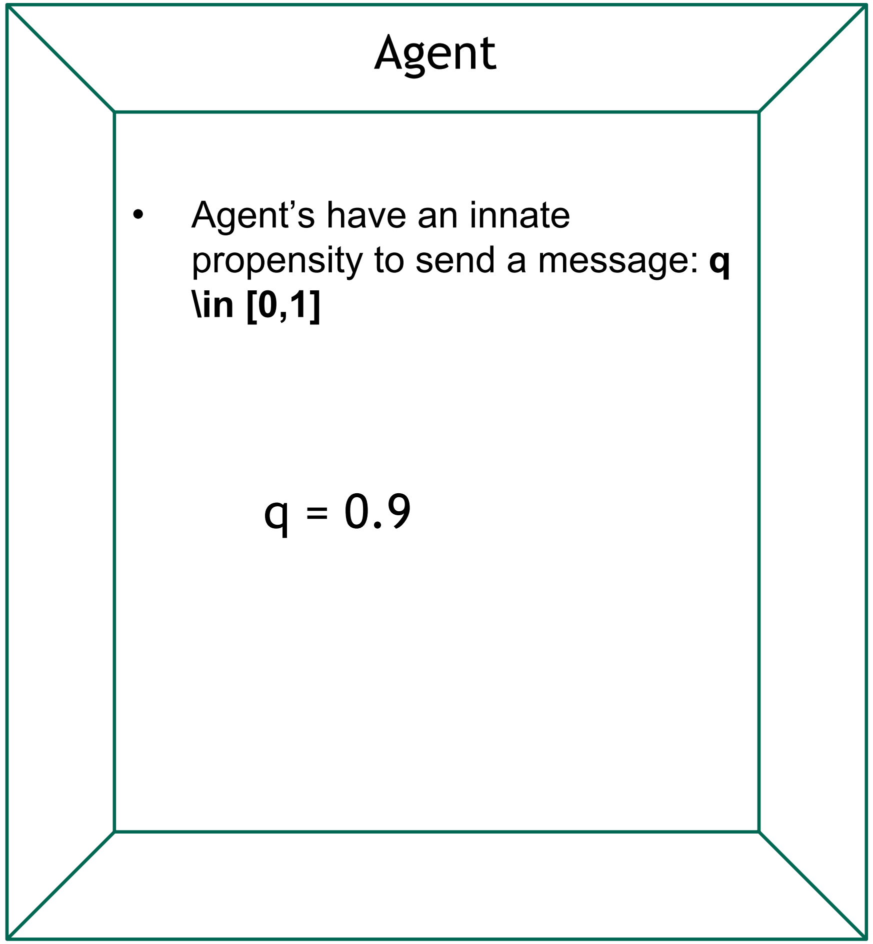
- Agent-based stochastic model.
- Captures:
  - Individual behavior and attention.
  - Social network.
  - Information cascades.

INDIVIDUAL AGENT

Inbox

	11	1	10	15	8
ϕ	0.1	0.9	0.2	0.3	0.1

- **Inbox** includes all the messages received by the agent.
- Messages have an unique id, and innate “virality” measure \phi \in [0,1].
    - For now virality is determined per message, but you can imagine a situation where virality changes as time increases (rewarding novelty)

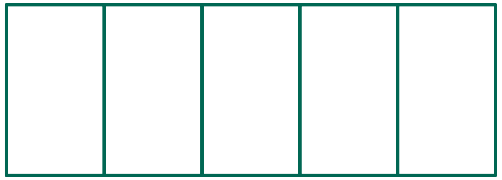


- Agent’s have an innate propensity to send a message: **q \in [0,1]**

q = 0.9

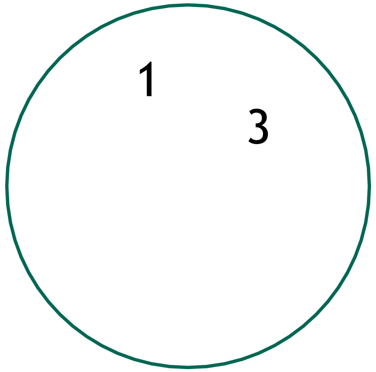
- **Outbox** are messages schedules to be sent this time point.

Outbox



- **Sent** are messages that have been sent by the agent.
- An agent can send a specific message only once.

Sent



INDIVIDUAL AGENT

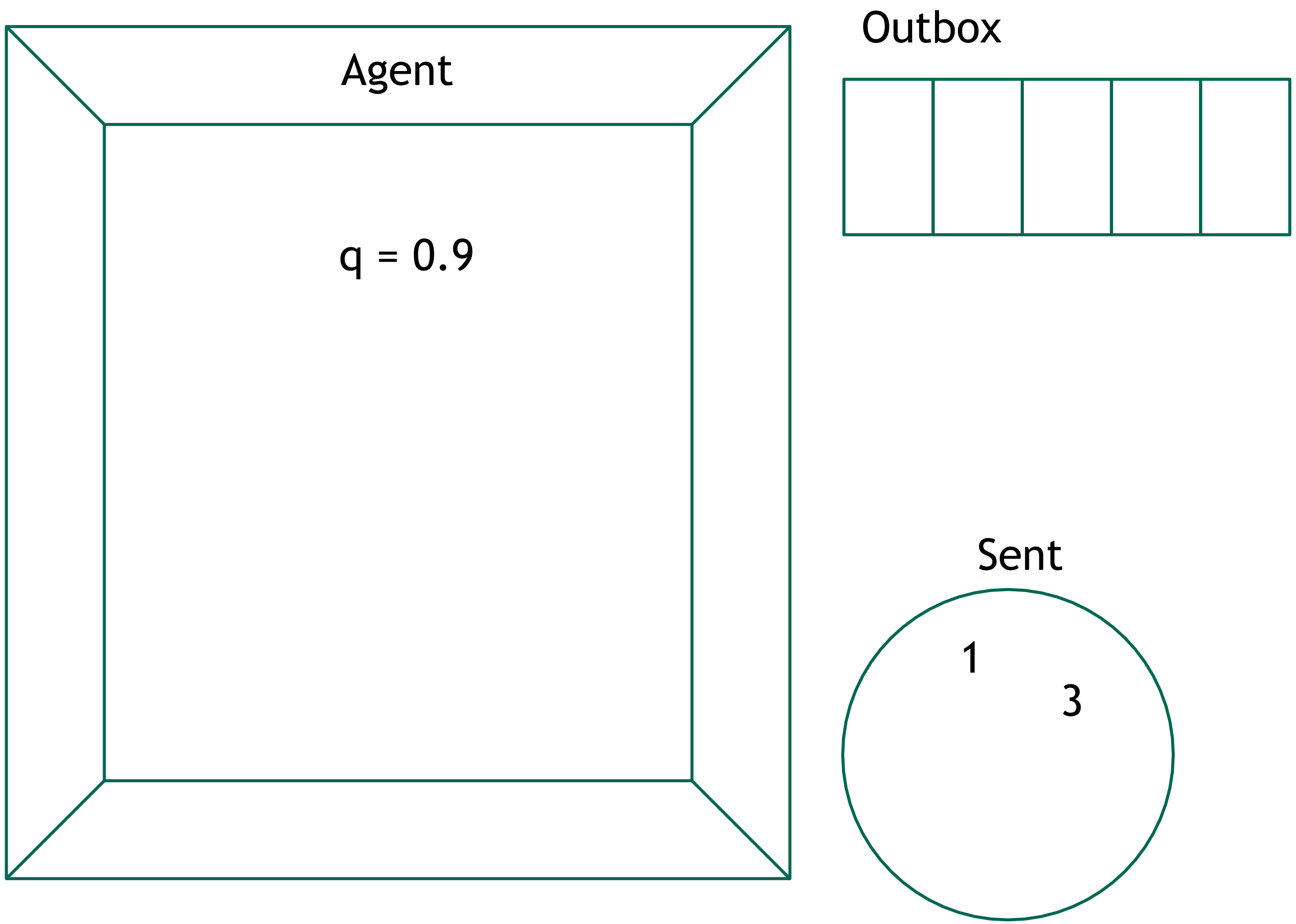
Time: 1

Inbox

ϕ					
	11	1	10	15	8
	0.1	0.9	0.2	0.3	0.1

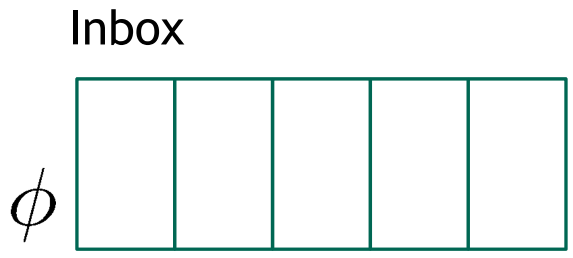
K = 3

- K messages are randomly chosen each time point.
  - Represents cognitive constraints and informational overload.
- Remaining messages in inbox are discarded.

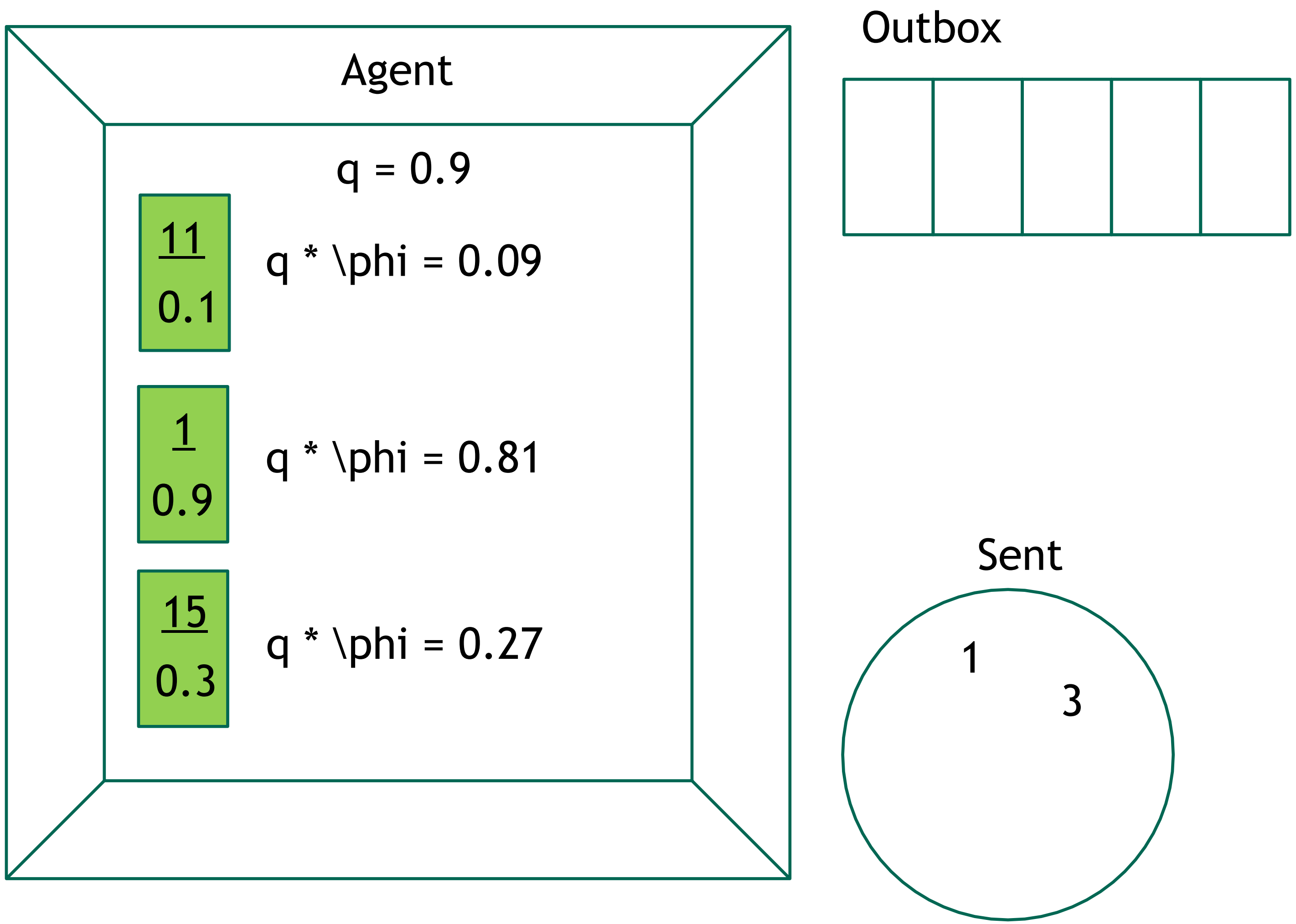


INDIVIDUAL AGENT

Time: 1



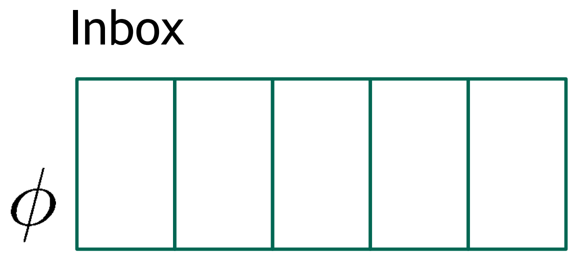
K = 3



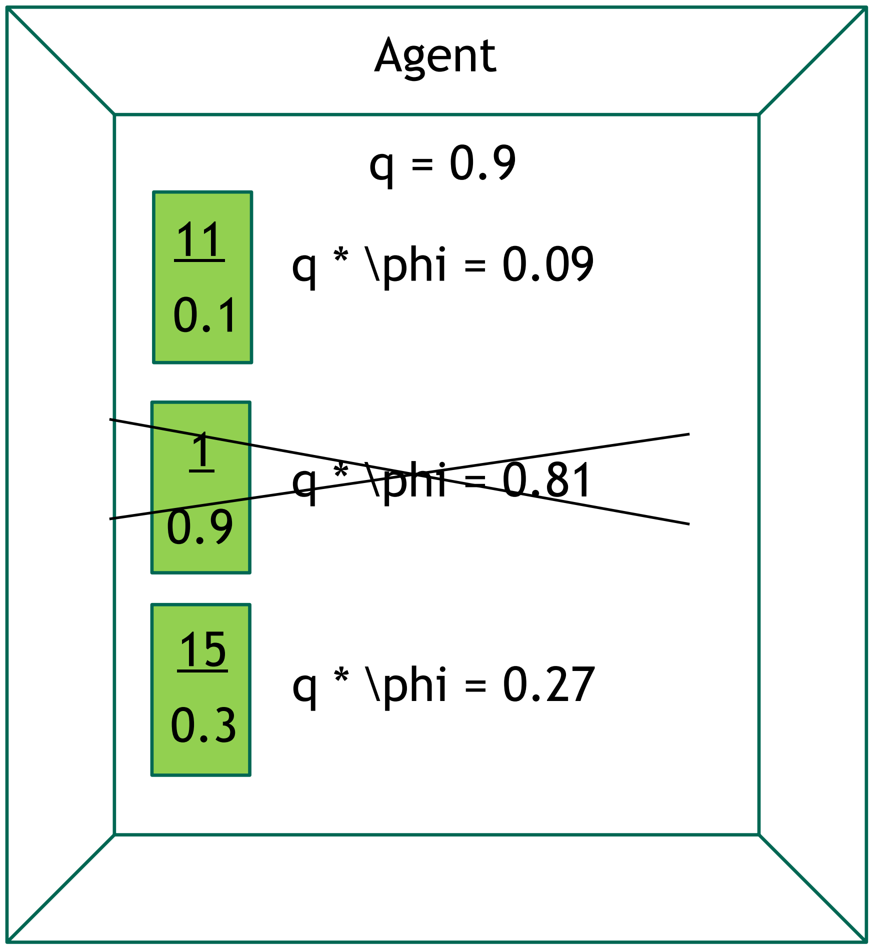
- Each messages is evaluated to see if it should be resent.
- Current: P(resend message) = q* \phi.
- This needs to change though.

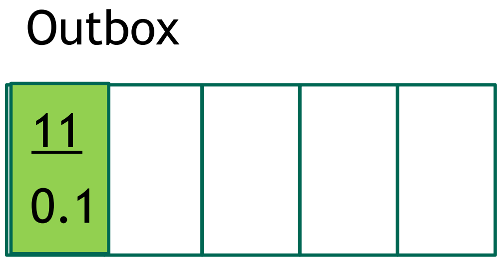
INDIVIDUAL AGENT

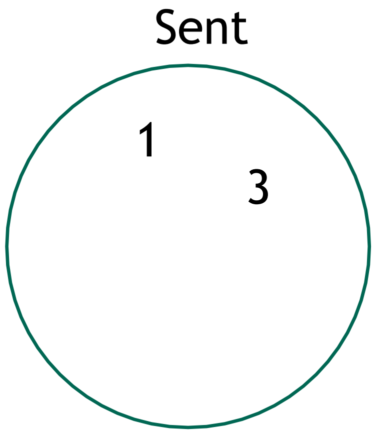
Time: 1



K = 3



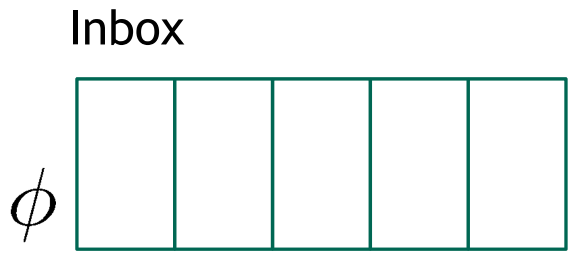




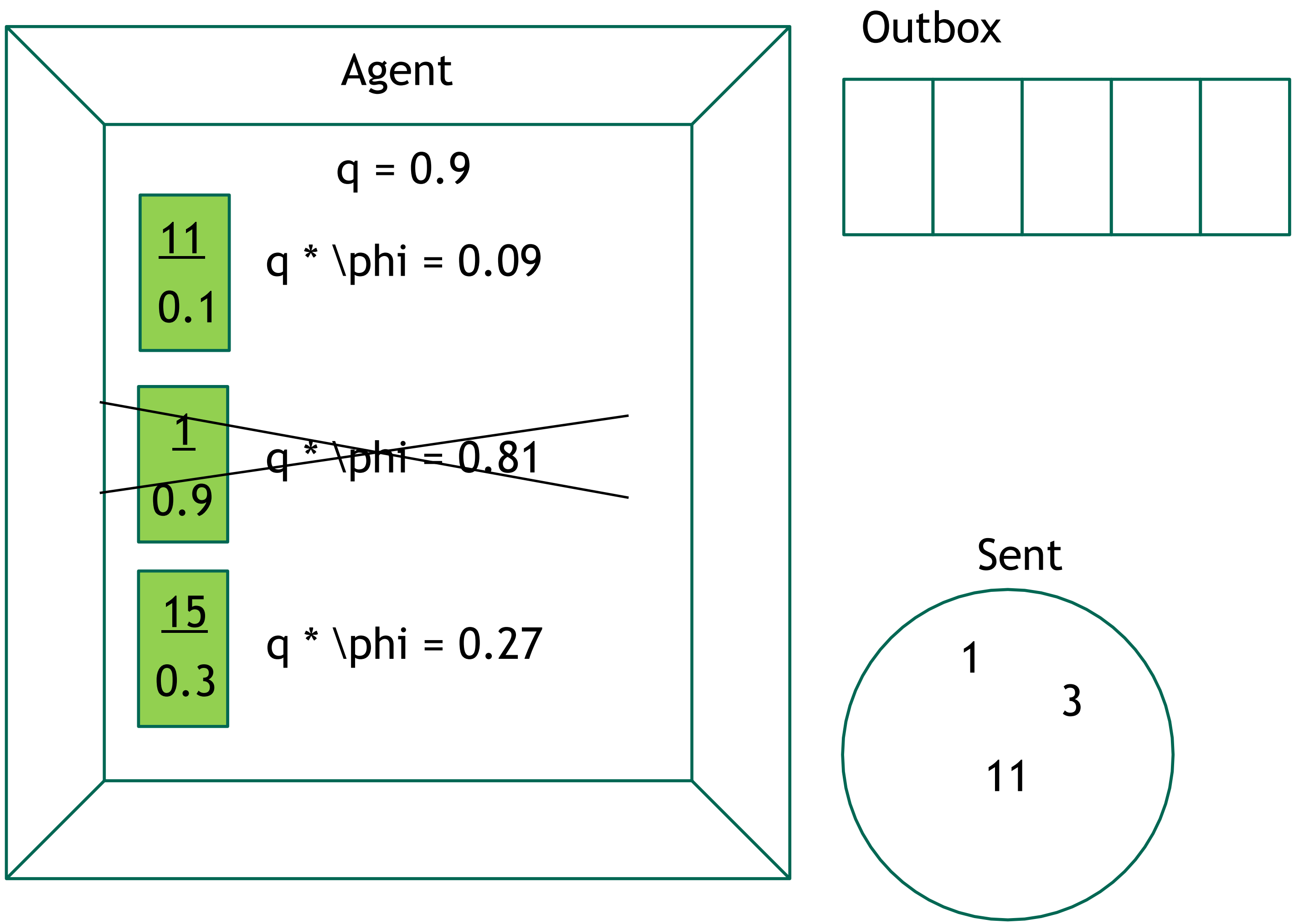
- Messages that have been sent before are discarded.
- Pick messages based on probability of resending.

INDIVIDUAL AGENT

Time: 1



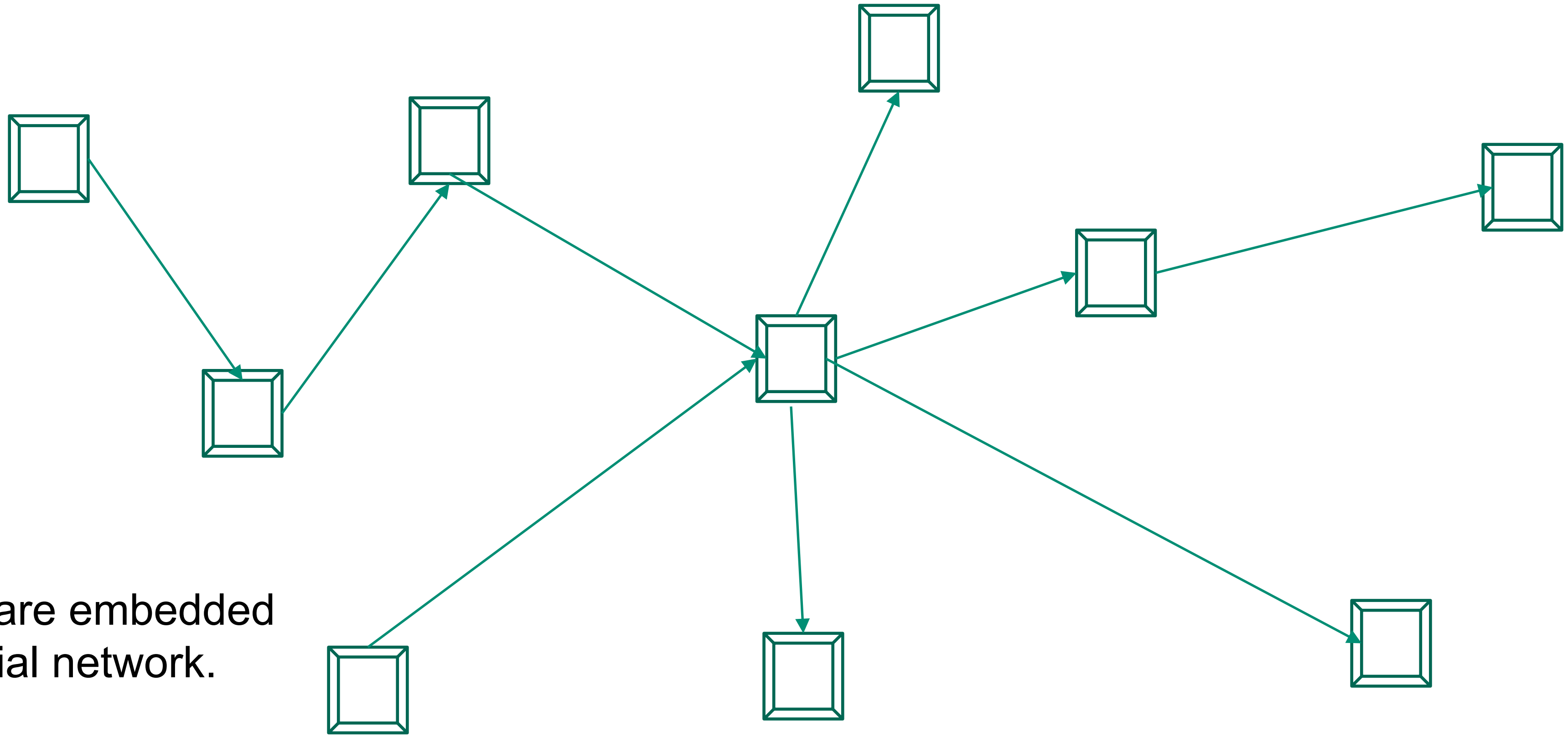
K = 3



- Messages from outbox will be distributed to all neighboring agents.

POPULATION LEVEL

Time: 1



- Agents are embedded in a social network.

SIMPLE INFORMATION DIFFUSION MODEL

- Agent-based stochastic model.
- Captures:
  - Individual behavior and attention.
  - Social network.
  - Information cascades.
- Parameters:
  - Social network topology
  - q_i – innate propensity to send a message.
  - k_i -- Number of messages seen by the agent.
  - Message seeding:
    - Number of messages seeded in simulation every Nth tick with virality.

NETWORK PARAMETERS

	Density		
	Low	Medium	High
Random (ER)	P = 0.006	P = 0.01	P = 0.02
Scale Free (SF)	M = 3	M = 5	M = 10
Small World (SW)	K = 6	K = 10	K = 20
(small world-y)	P = 0.1	P = 0.2	P = 0.4

AGENT PARAMETERS

	Low	Medium	High
Q_i			
mean	0.8	1.0	1.2
sd	5% of mean	20% of mean	50% of mean
K_i			
mean	1	5	10
sd	5% of mean	20% of mean	50% of mean

MESSAGE PARAMETERS

	Low	Medium	High
Number of messages			
seeded at start added every nth tick	100	250	500
Number of seeded agents	1	10	25
Nth tick	1	10	25
Virality power	5	20	50

RELATIVE ERROR

$$\omega_{i,j} = \frac{Predict(M_i, P_j^{val})}{Predict(M_i, P_i^{val})}$$

- Measures how well M_i performs on a new “world” based on trained world.
- Varies between $[0, +\infty]$
  - If $\omega_{i,j} < 1.0$ then M_i performs better on the new data than the data it was trained on.
  - If $\omega_{i,j} \geq 1.0$ then M_i performs worse on the new data than the data it was trained on.

THE INTUITION CHECK WORKS

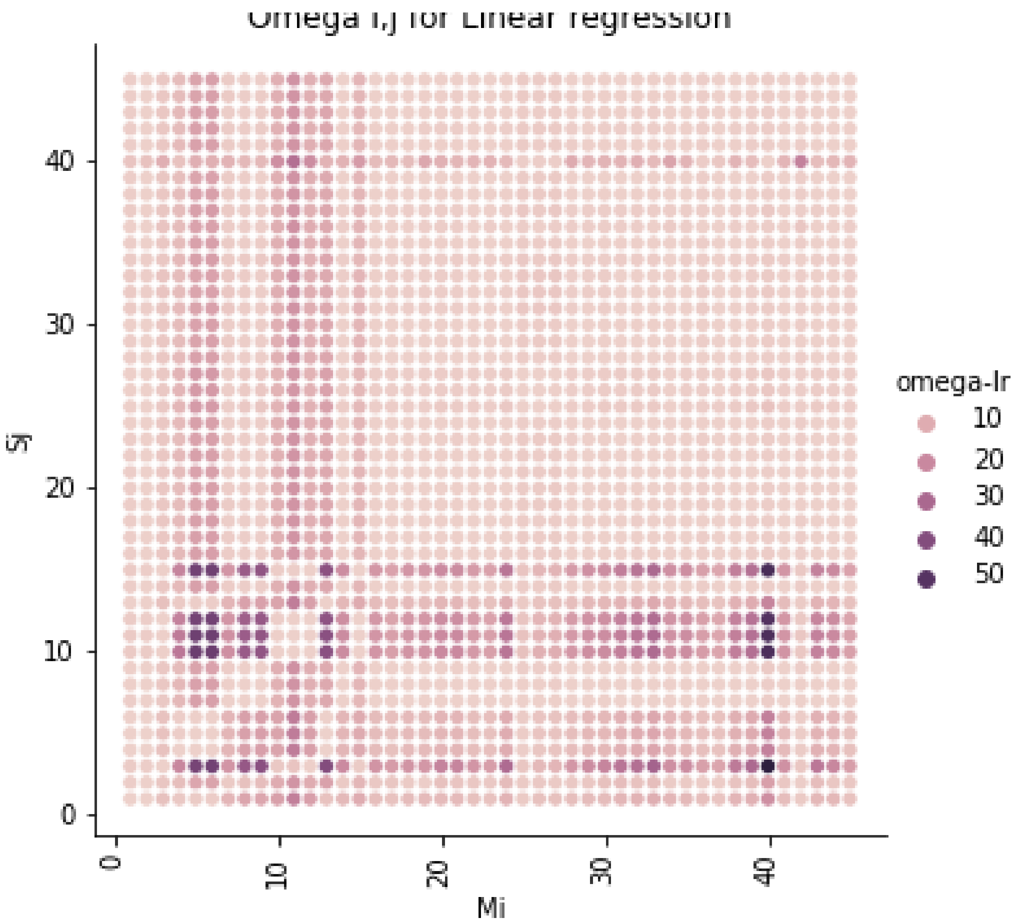
- Intuitively: when the methods are trained on a simulation that is the same as which they are predicting, their error should be lower then when they train and predict on differing simulations.
  - Underlying patterns are not changing.
- Compare $\mathrm{mean}(Predict(M_i,S_i))$ against $mean\left(Predict\left(M_i,S_j\right)\right), \forall i \neq j$ Data used is the validation set from each simulation.

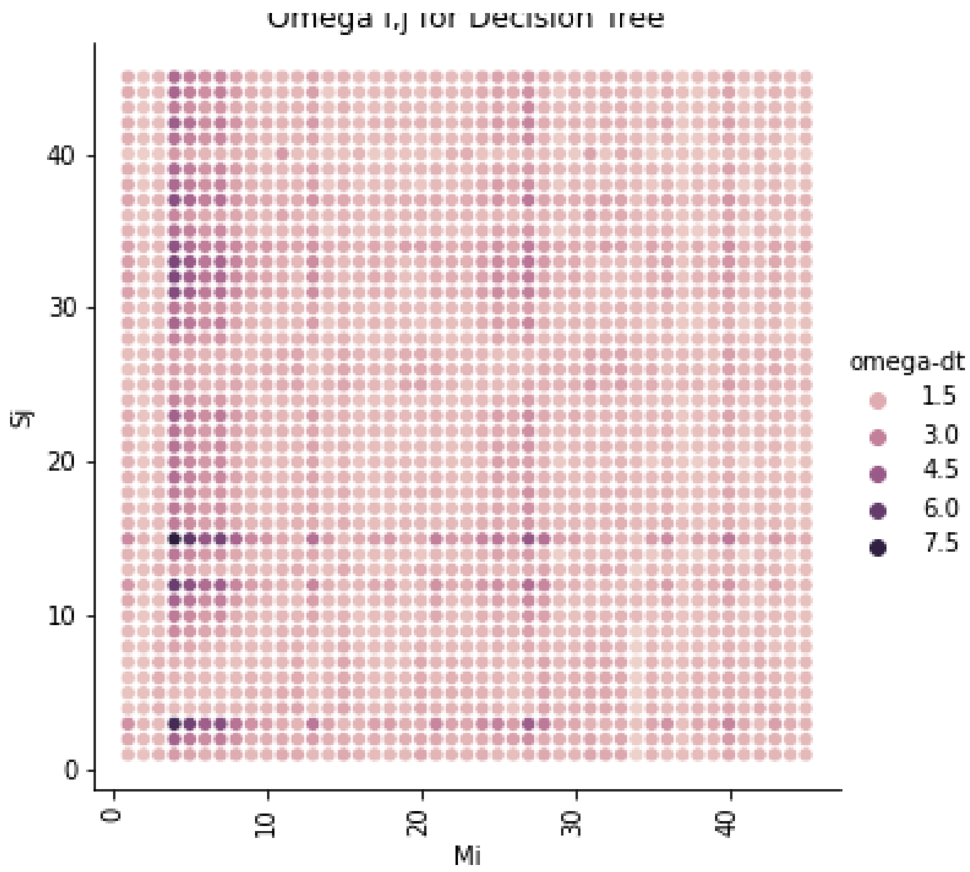
	RMSE LR		RMSE DT	
	mean	std	mean	std
sameParamTT				
False	676.353675	949.270816	176.641065	81.256355
True	131.296880	52.301123	143.730372	84.549962

	Omega LR		RMSE DT	
	mean	std	mean	std
sameParamTT				
False	5.443542	7.398412	1.381543	0.713567
True	1.000000	0.000000	1.000000	0.000000

- We see RMSE is lower (on average) for $Predict(M_i,S_i)$.
- Decision tree does better when predicting over other simulations it wasn't trained on.

THE SIMULATION YOU TRAIN ON MATTERS.





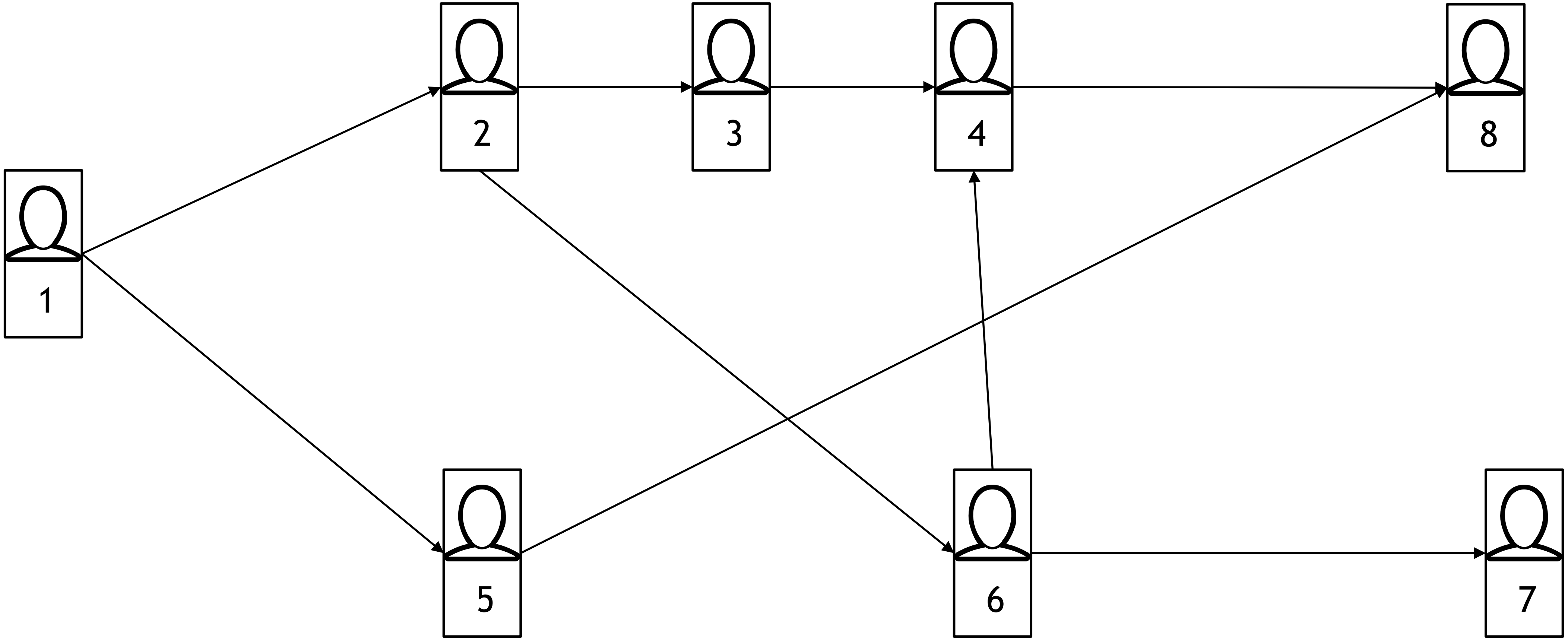
DECISION TREE SEEMS MORE ROBUST TO DATASET SHIFT

- Max value of $\omega_{i,j}$ is very different between Linear Regression and Decision tree.
- Less variability in decision tree.
- Some simulations are better to train against than others.
  - Training on simulation 5 resulted in high mean ω for both linear regression and decision tree.
- We would argue that decision trees are more “robust” to changes in dataset shift.

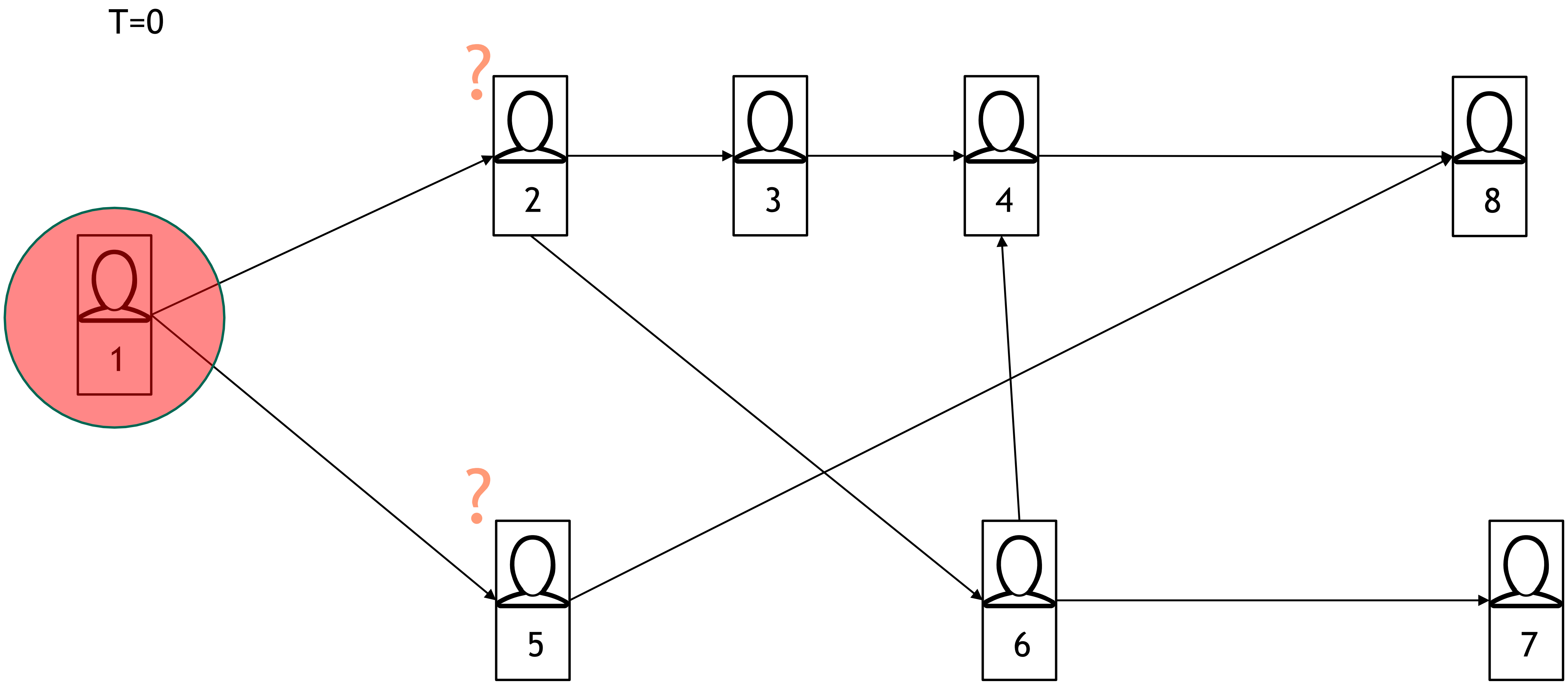
NEXT STEPS

- Developing a more complex model of information diffusion that integrates:
  - Ideological consistency.
  - Trust/Centrality.
  - Information accuracy.
- Expanding the data analytic methods to include Deep Neural Network.
- Comparing model output to real world-data.

DATA ANALYTICS QUESTIONS

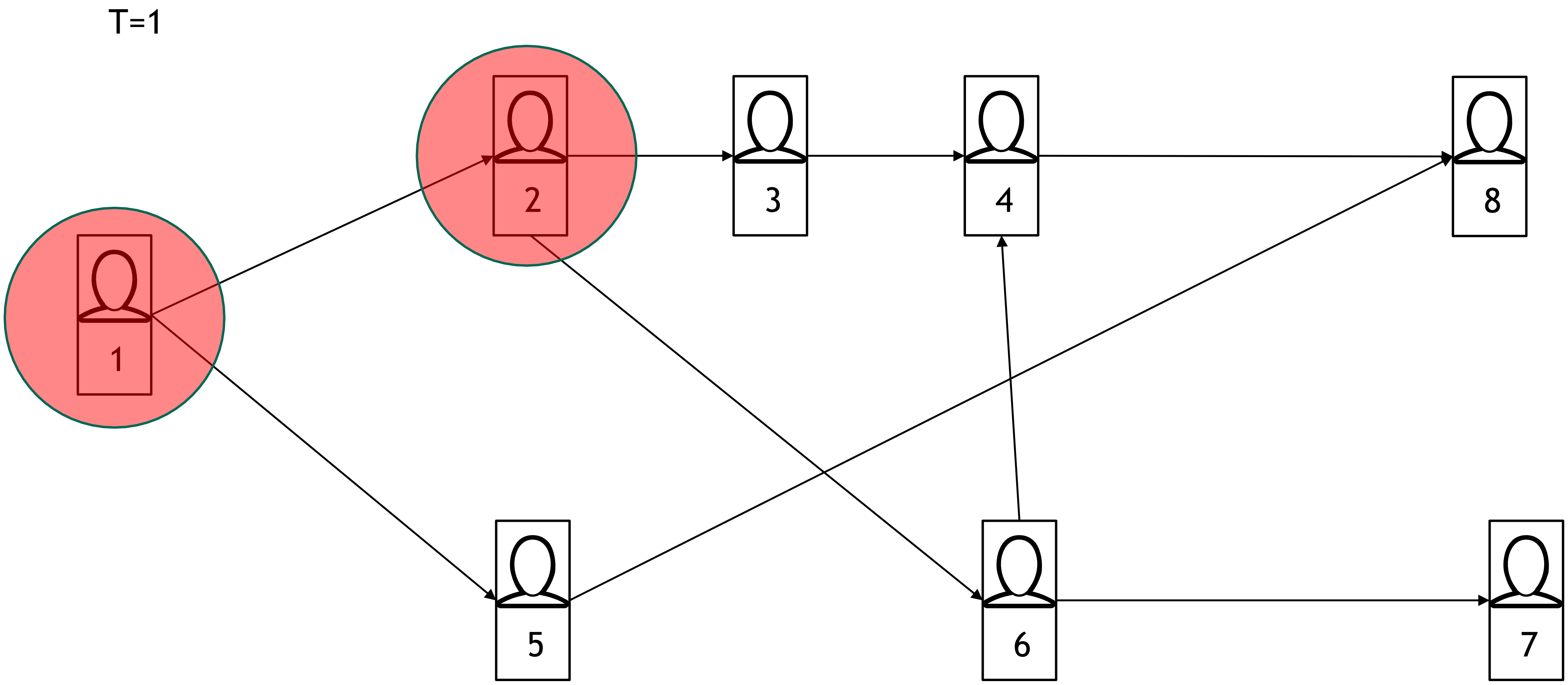


CASCADE OF INFORMATION



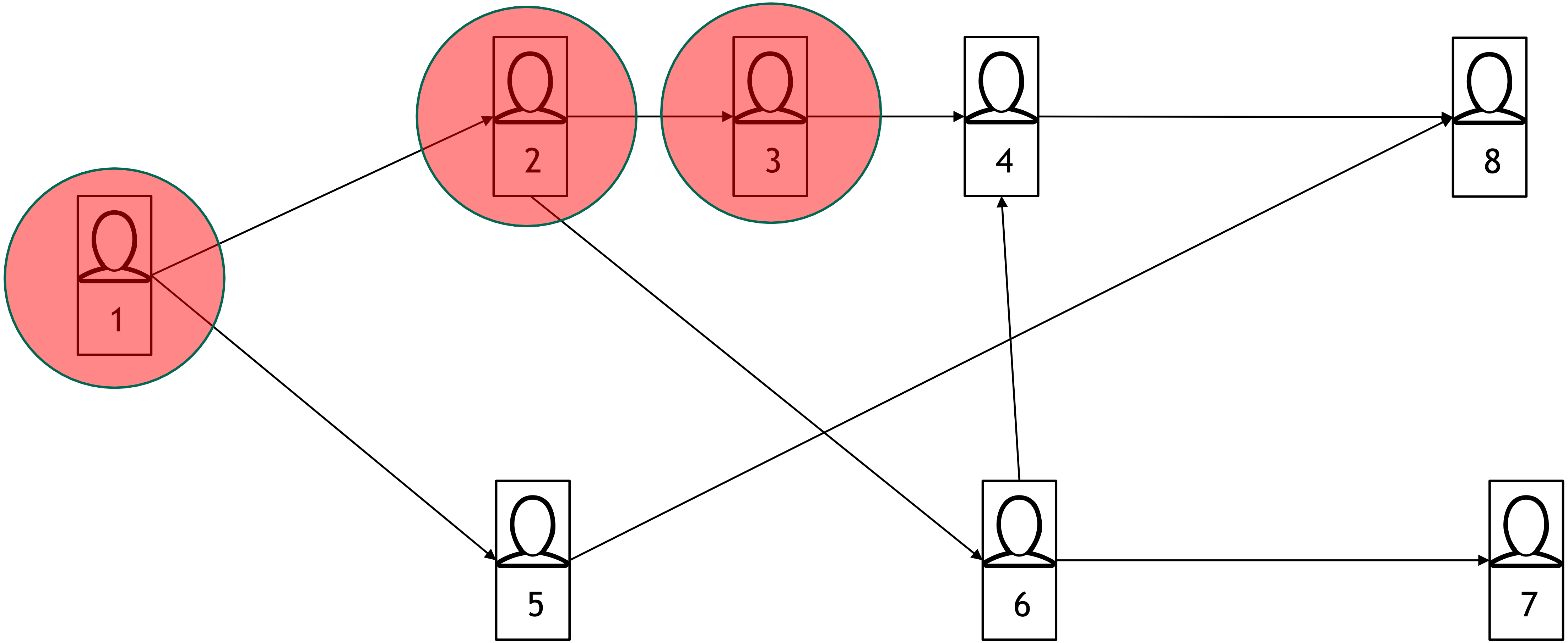
Micro prediction task: Predict next infected user.

CASCADE OF INFORMATION

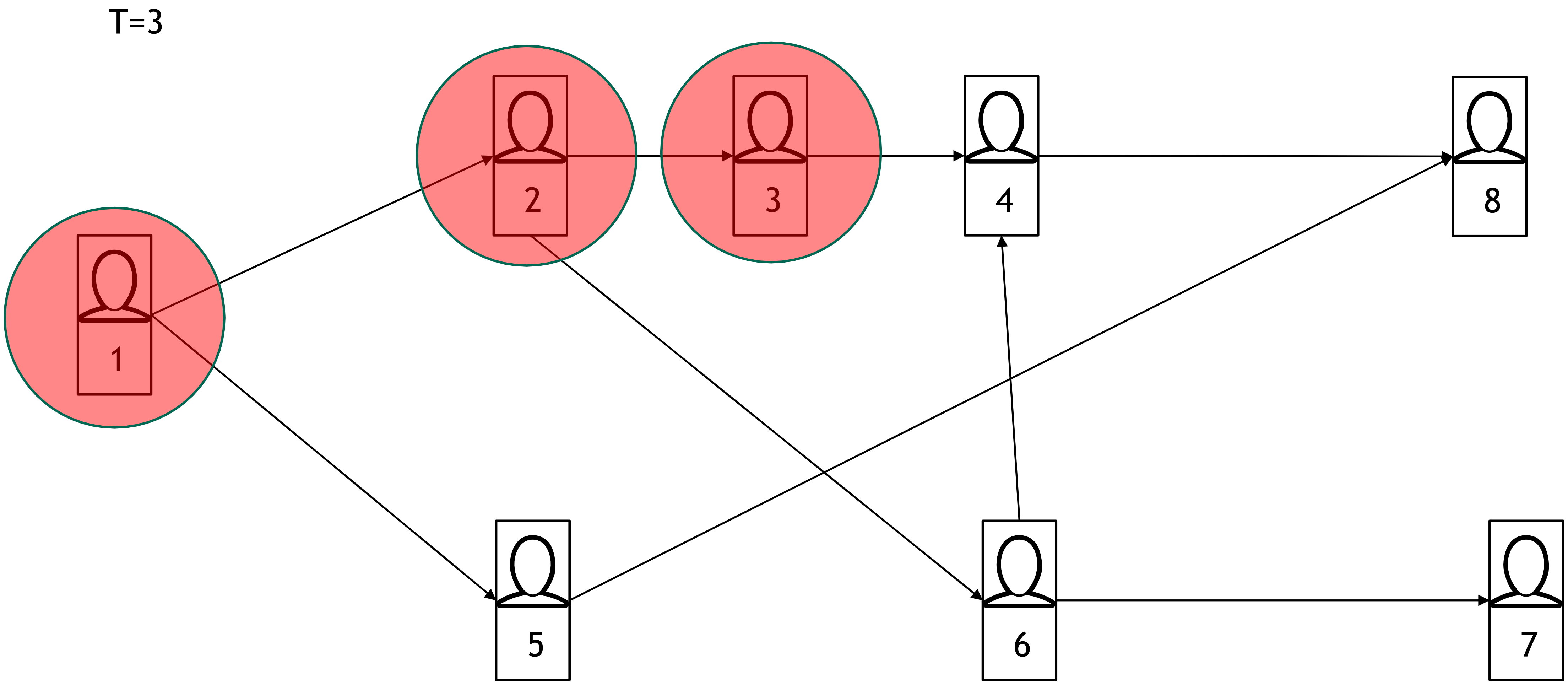


CASCADE OF INFORMATION

T=2



CASCADE OF INFORMATION

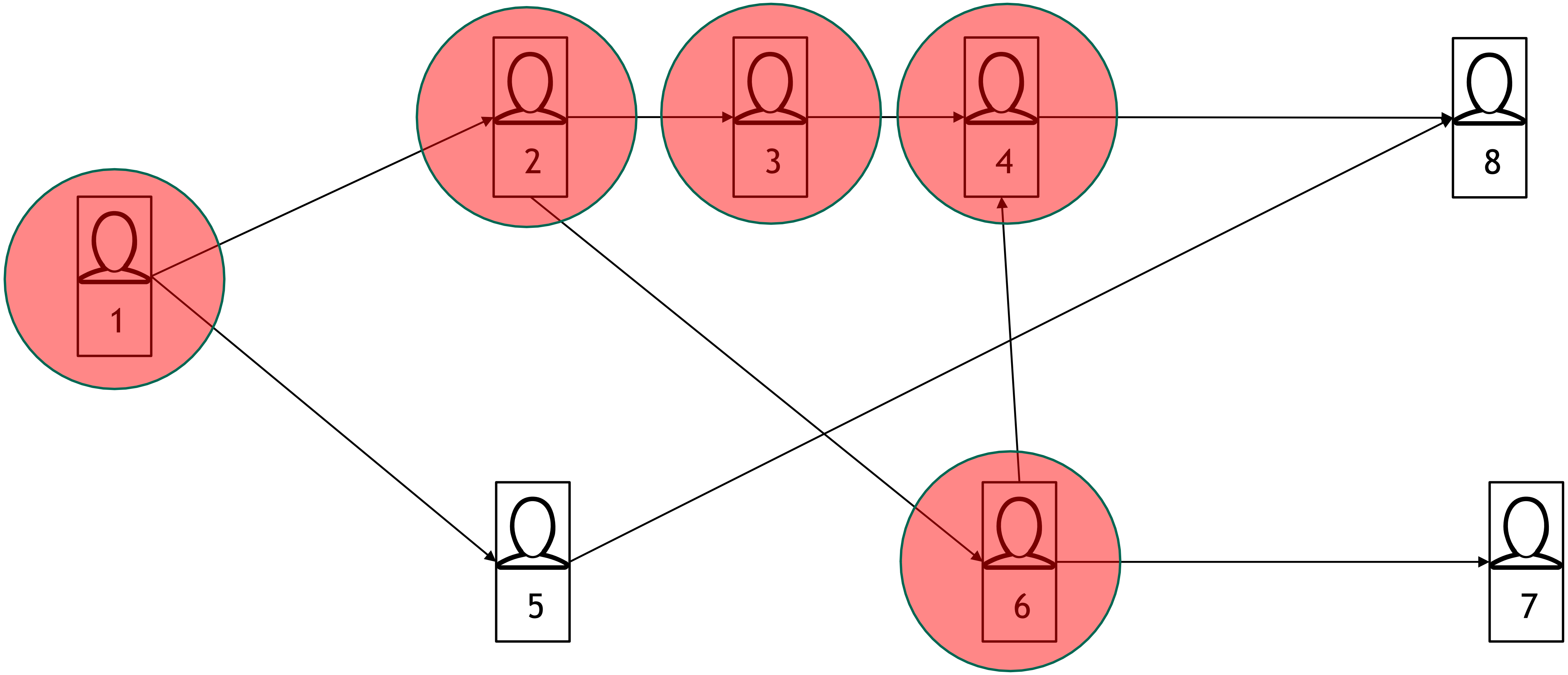


CASCADE OF INFORMATION

T=4

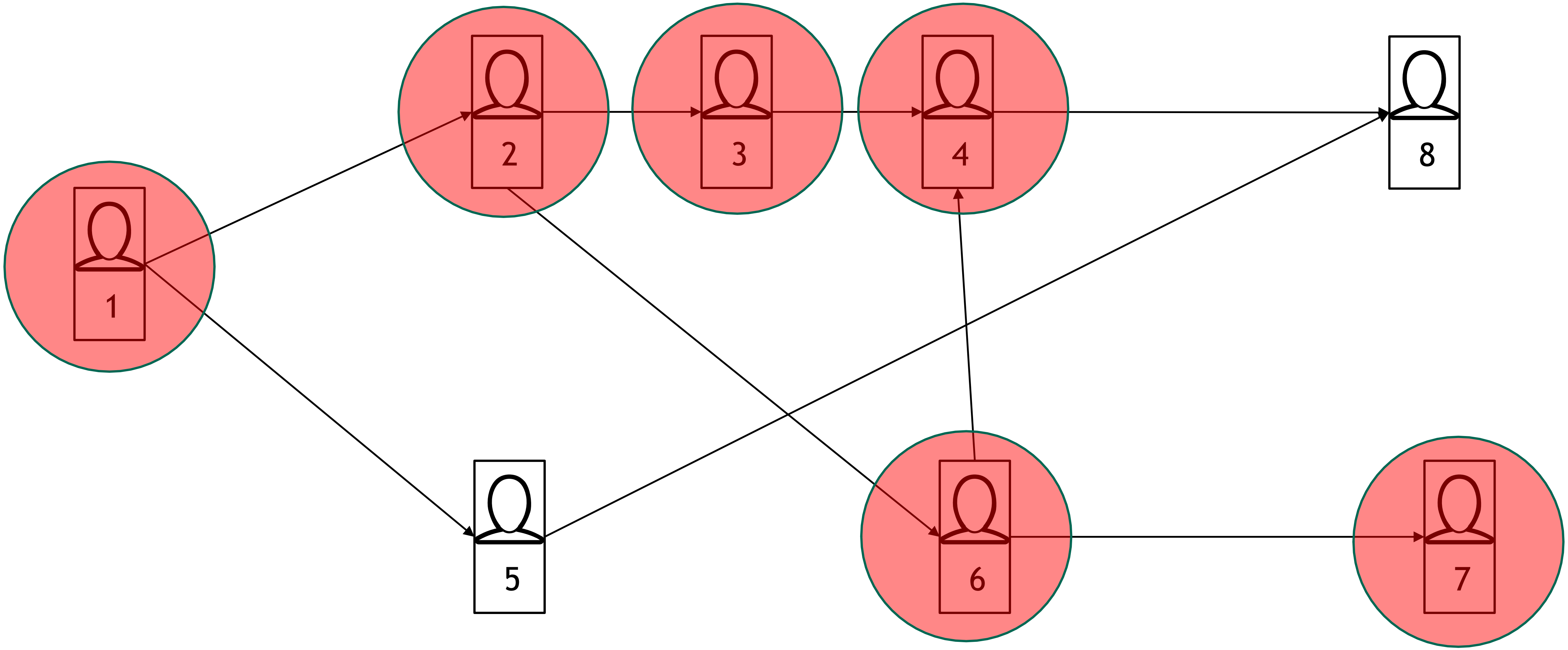
Note multiple people adopt the disinformation.

Was node 4 influenced by node 3? Time delay for node 6.



CASCADE OF INFORMATION

T=4
Note multiple people adopt the disinformation.
Was node 4 influenced by node 3? Time delay for node 6.



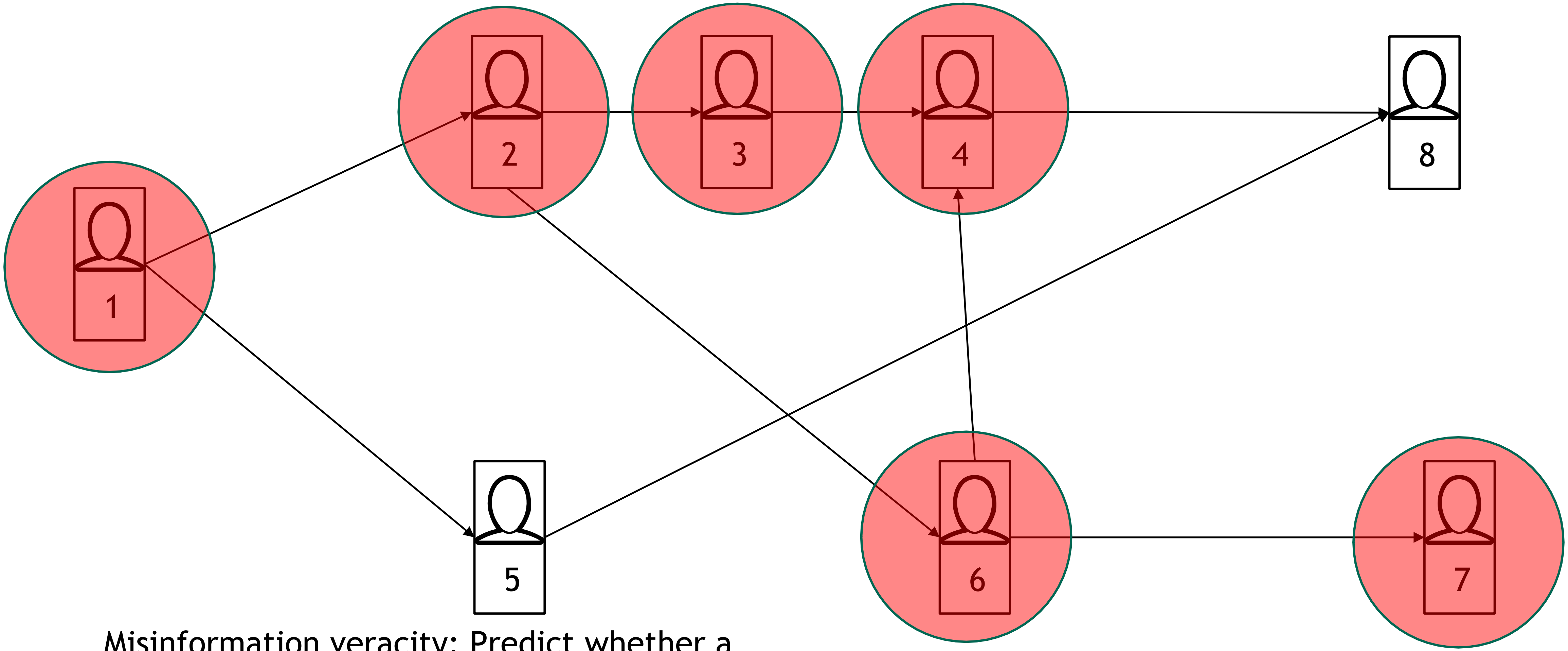
Macro prediction task: Predict size of the cascade (6).

CASCADE OF INFORMATION

T=4

Note multiple people adopt the disinformation.

Was node 4 influenced by node 3? Time delay for node 6.



Misinformation veracity: Predict whether a cascade contains misinformation or not based on structure of the cascade.