Social Reinforcement Learning to Combat Fake News Spread

Mahak Goindani

Department of Computer Science Purdue University,
West Lafayette, IN mgoindan@purdue.edu

Jennifer Neville

Departments of Computer Science and Statistics
Purdue University,
West Lafayette, IN neville@purdue.edu

Presented by

Rick Rejeleene
PhD Information Science

Outline

- Background and Context of Reinforcement Learning
- Exploration of the Research Paper

Background of Reinforcement Learning

- Reinforcement Learning is an area of Machine Learning
- It's concerned about how a software agent ought to take actions in an environment to maximize a cumulative reward
- It is the science of decision making

Background of Reinforcement Learning

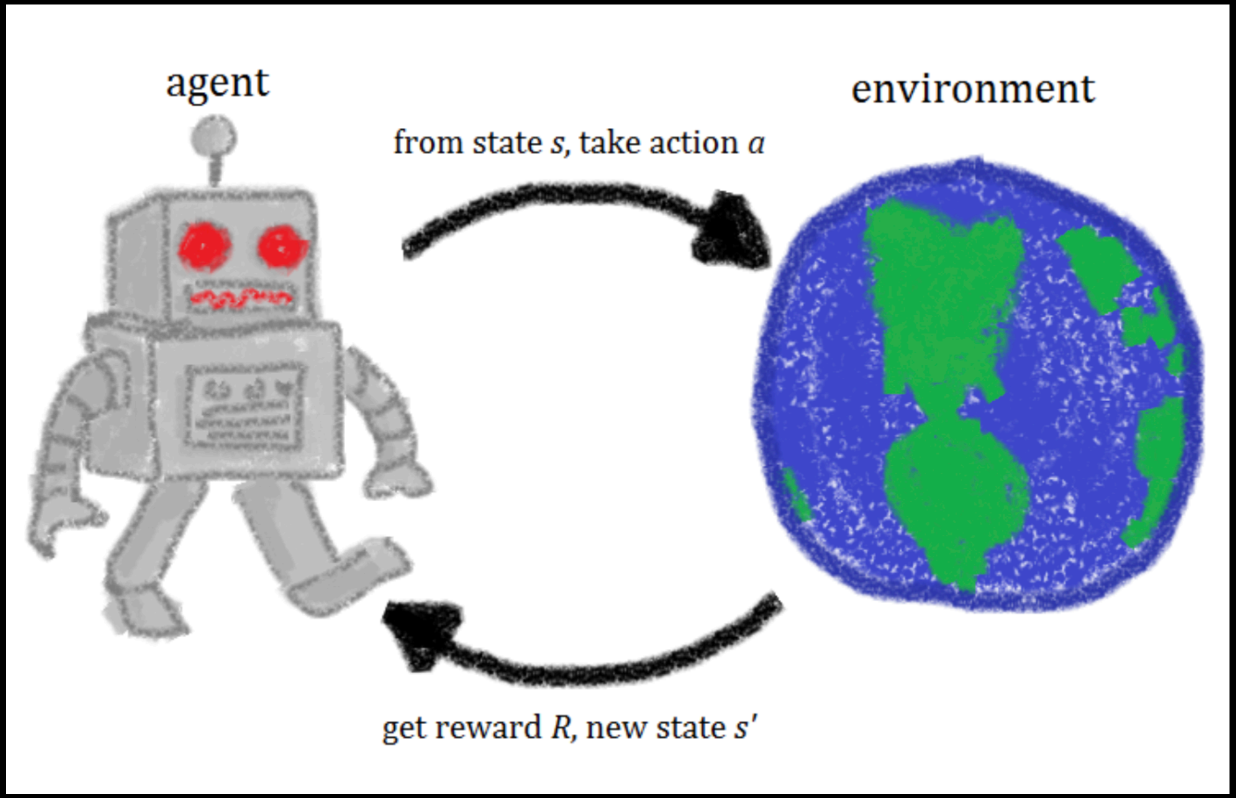
- It’s studied in Game theory, Operations Research, Information theory, statistics
- A basic Reinforcement is modeled as Markov Decision Process

What does it mean to solve a RL problem?

-Solving a reinforcement learning task means, roughly, finding a policy that achieves a lot of reward over the long run

- It involves problems learning what to do, how to map situations to actions to maximize numerical signal reward

How does it Work?



- All Reinforcement Learning methods have the below three in common
- Estimation of Value Function
- All Operate by backing up values along actual or possible state trajectories
- Generalized Policy Iteration

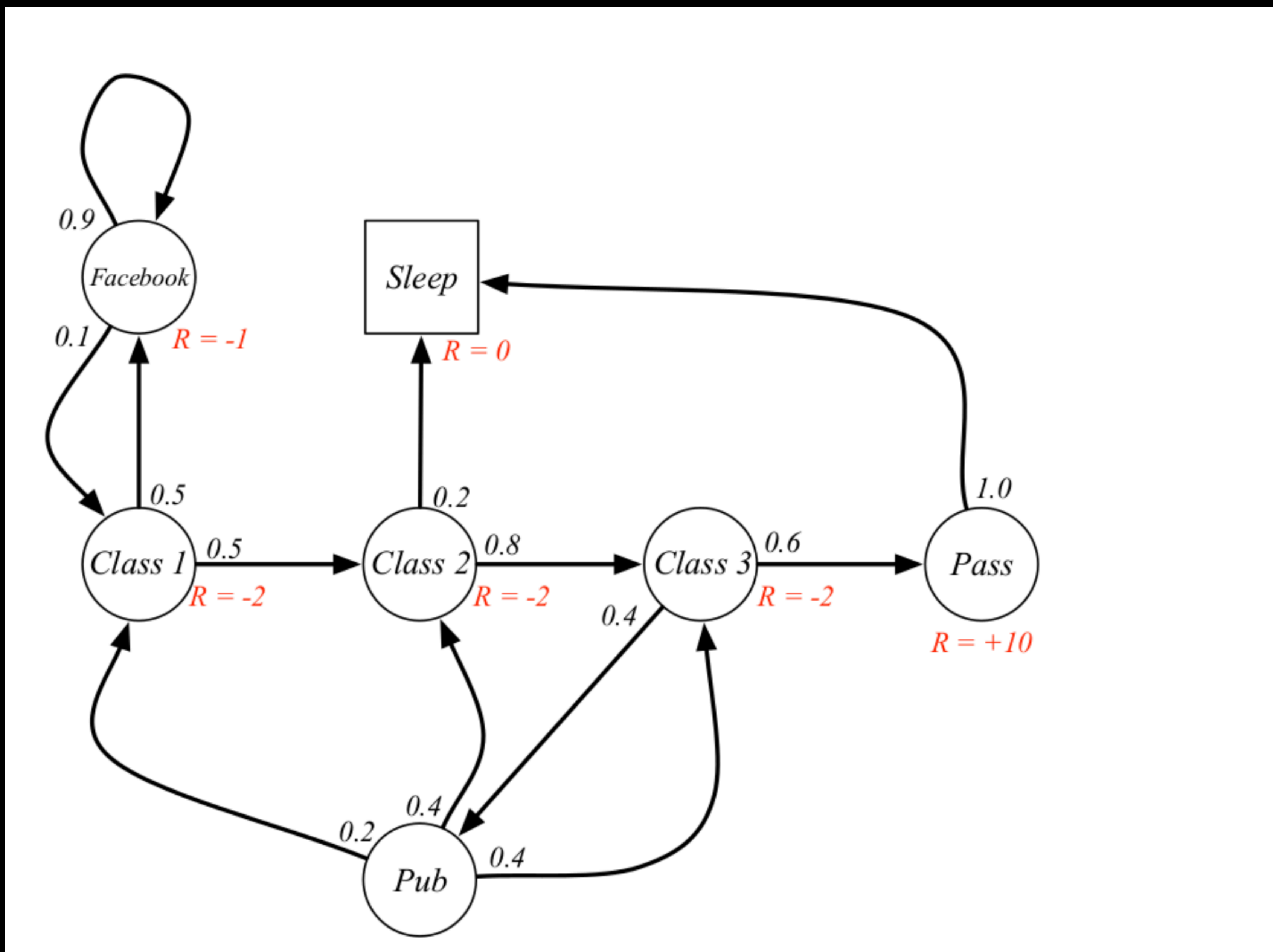
How does it work?

Basic reinforcement is modeled as a Markov decision process:

- a set of environment and agent states, S;
- a set of actions, A, of the agent;
- $P_a(s, s') = \Pr(s_{t+1} = s' \mid s_t = s, a_t = a)$ is the probability of transition from state s to state s' under action a.
- $R_a(s, s')$ is the immediate reward after transition from s to s' with action a.
- rules that describe what the agent observes

- Formalization of Reinforcement Learning

Markov Decision Process

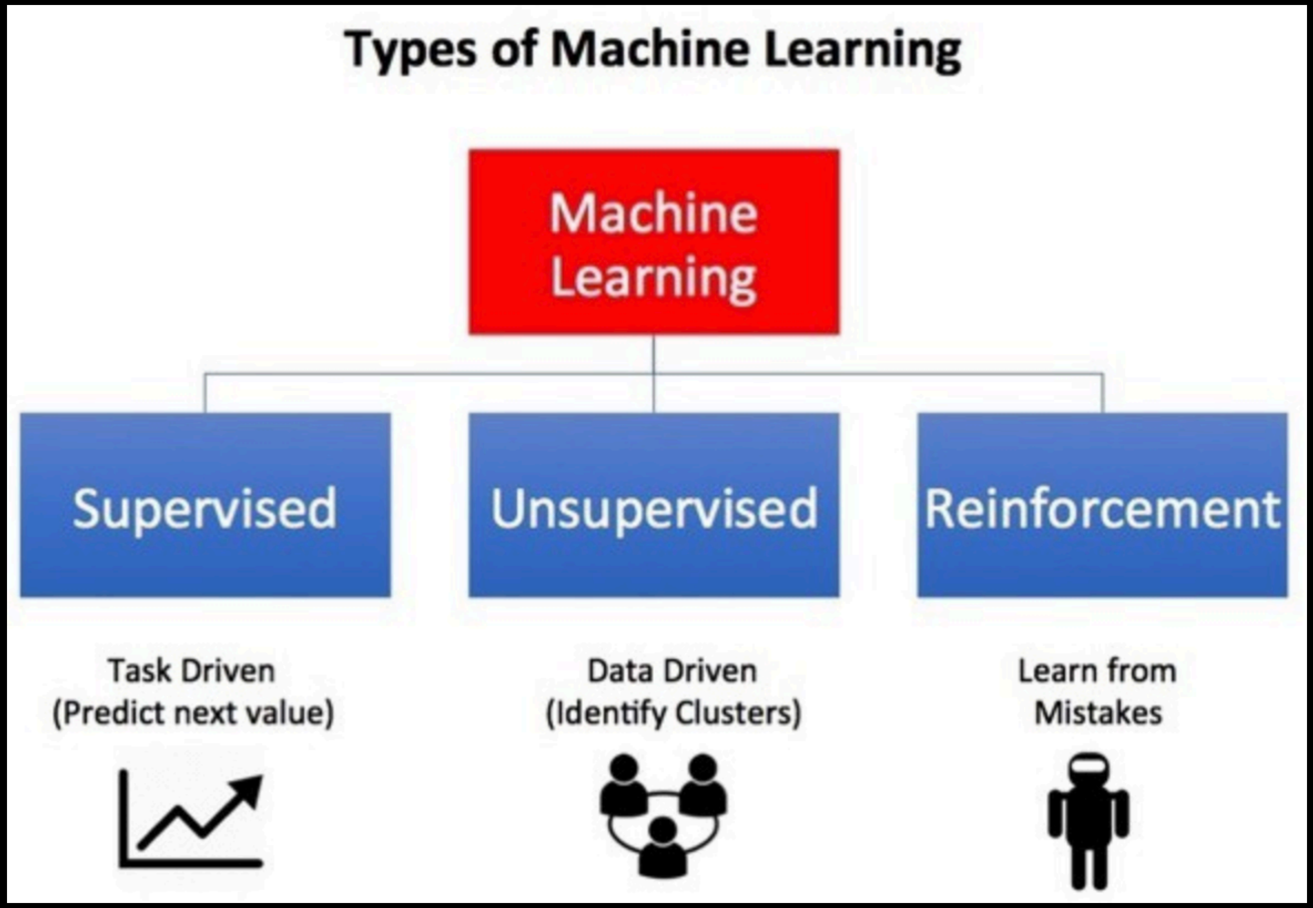


*https://cdn-images-1.medium.com/max/1600/1*yTz8cG5yfov_kdjKz6dNJg.png*

- Reward at the end - Pass
- States: Class 1, Class 2, Sleep, Facebook, Sleep,
- Probability of each state
- Sleep is terminal state
- Markov Reward Process is R
- R is how much reward accumulated through a sequence

Where does it Belong?

- Broadly under the field of Artificial Intelligence
- Specifically under Machine Learning

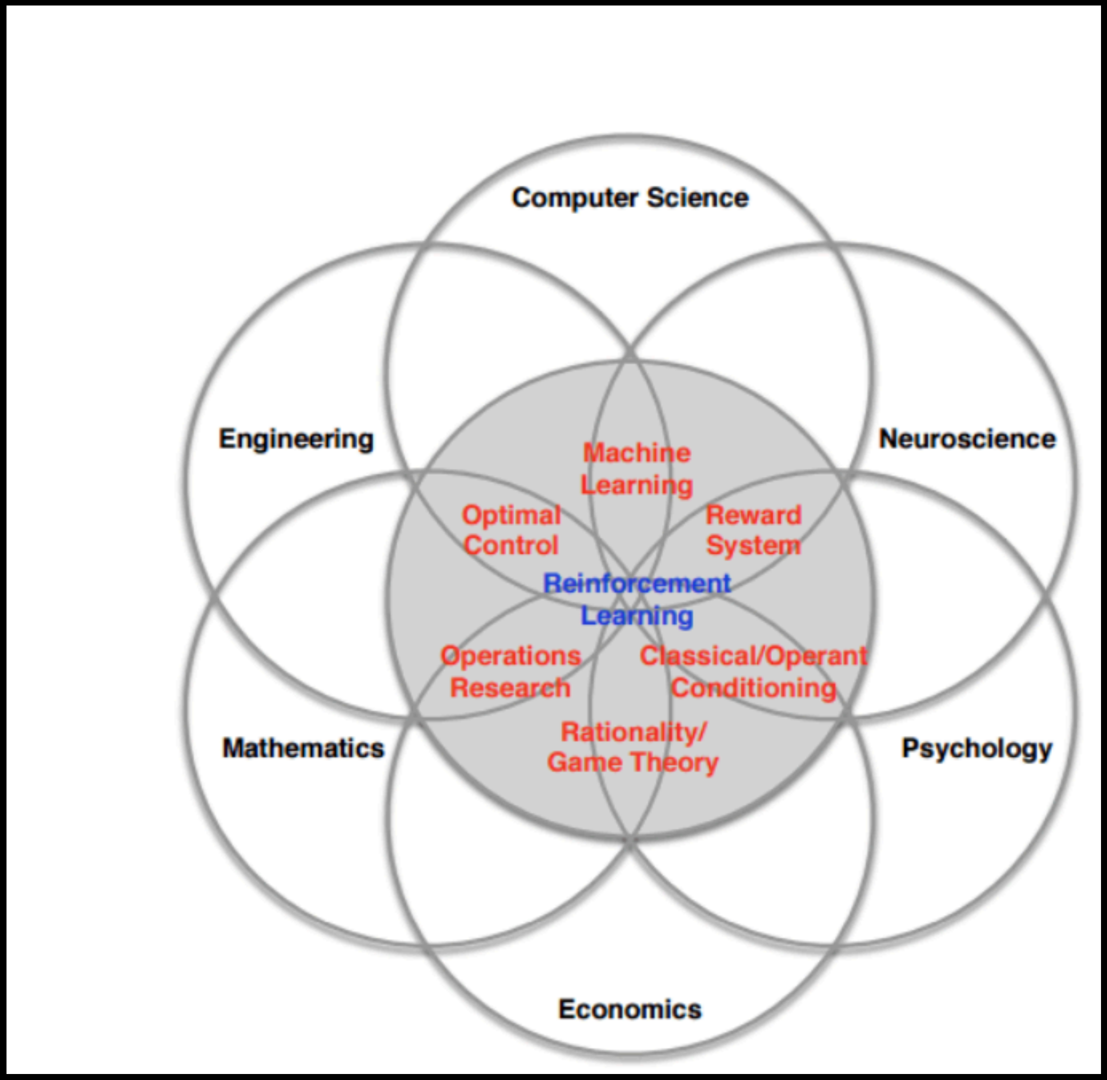


http://en.proft.me/media/science/ml_types.png

What’s the difference between unsupervised vs Reinforcement Learning?

- Unsupervised Learning is about finding hidden pattern
- Reinforcement Learning is finding maximum optimal output

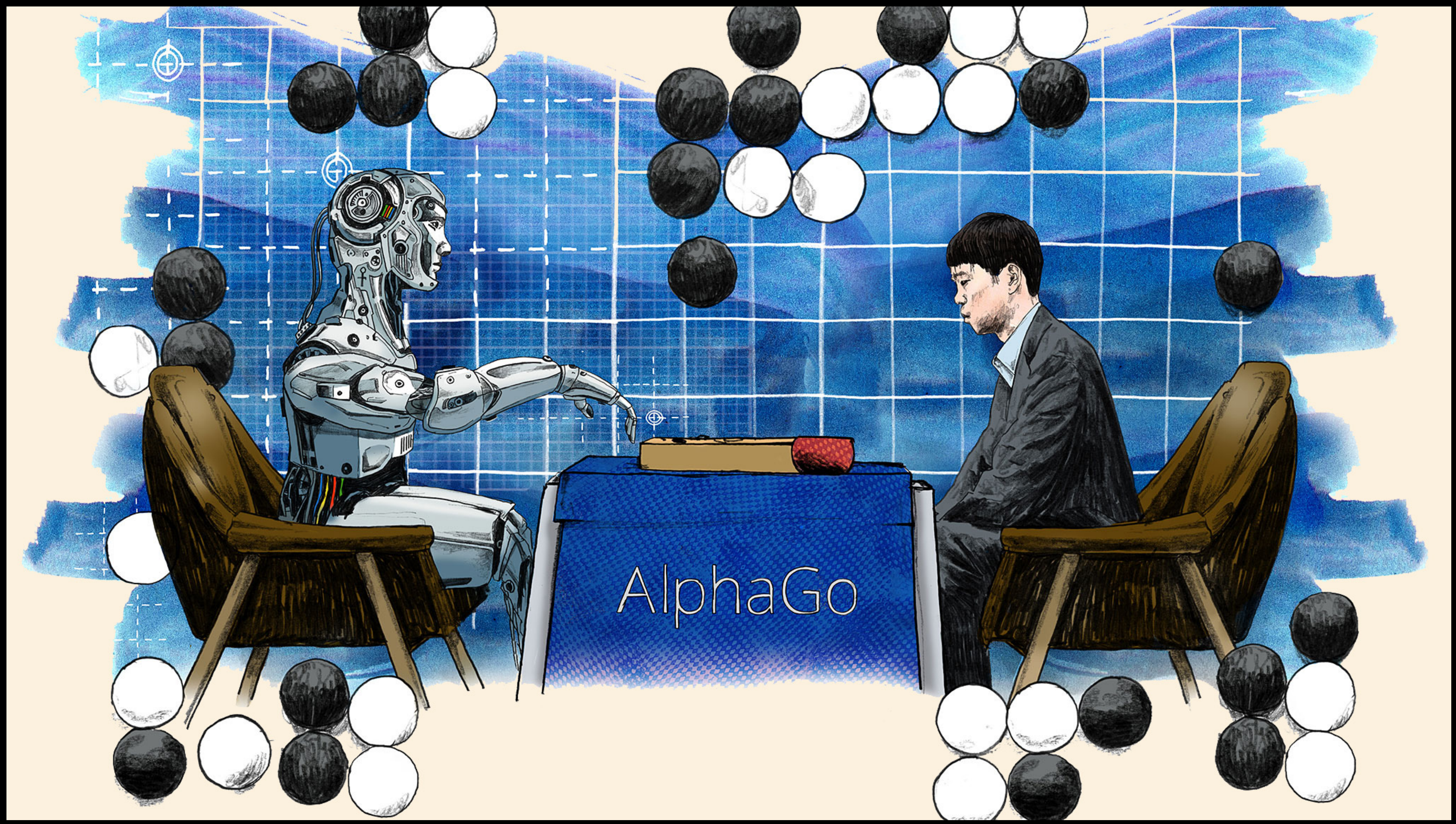
Why Reinforcement Learning



https://i1.wp.com/syncedreview.com/wp-content/uploads/2017/02/5.png?fit=768%2C613&ssl=1

- Cutting edge research
- Interdisciplinary research

Why Reinforcement Learning?



https://www.ft.com/content/cada14c4-d366-11e6-b06b-680c49b4b4c0

- Google DeepMind's AlphaGo (2016)
- Google DeepMind - mastered the impossibly complex board game Go

Why Reinforcement Learning?

AlphaGo beats human Go champ in milestone for artificial intelligence

LATEST WORLD-NATION

WORLD & NATION

The people sing: 'Les Mis' soothes, breaks Hong Kong hearts

Dec. 11, 2019

WORLD & NATION

Undecided voters are key target on eve of British election

Dec. 11, 2019

WORLD & NATION

Jersey City mayor says gunmen targeted kosher market, stops short of calling attack anti-Semitic

Dec. 11, 2019

WORLD & NATION

Newsletter: Two impeachment articles, one trade deal

https://www.latimes.com/world/asia/la-fg-korea-alphago-20160312-story.html

- Google DeepMind, It beat one of the best human GO players in the world in a high-profile match (2016)

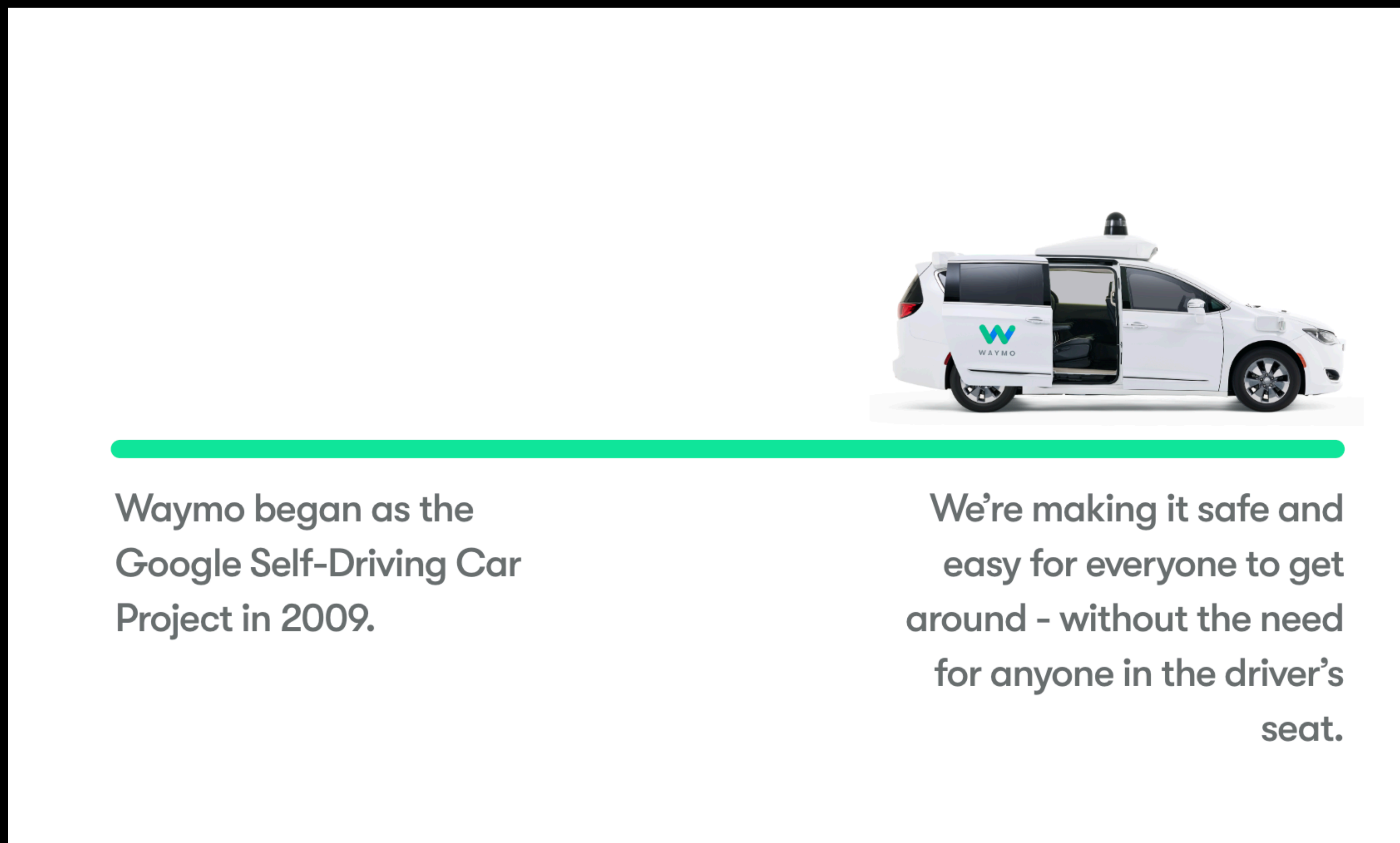
Why Reinforcement Learning?

Google AlphaGo

Why Reinforcement Learning?

- MIT Technology Review -10 Breakthrough Technologies 2017
- MIT Review - By Experimenting, Computers are figuring out how to do things that no programmer could teach them

Why Reinforcement Learning?



Waymo.com

- Self-Driving Cars: Already deployed on the road in Arizona
- WAYMO (Google), is beginning to offer members of the public rides in its self-driving cars in the Phoenix area.

Why Reinforcement Learning?

- Exciting new areas could be applied in the area of Social Media intersecting Reinforcement Learning
- Possible Areas:
    - Optimization News Recommendation
    - Decision Making in social bots, crawlers
    - Monitor fake news in Social Media
    - Maintaining

Applications of Reinforcement Learning

- Personalized Recommendation in Social Media
- Bidding and Advertisement
- Traffic Light Control
- Resources management in computer clusters
- Robotics
- Games
- Robotics
- Advertising
- Cardiology
- Resource Management
- Congestion Problem

Social Reinforcement Learning

- We are applying Reinforcement Learning in the context of Social Media through social networks
- We apply this approach to combat fake news

Key Idea of the Paper

"Social Reinforcement Learning approach to combat spread of Fake News"

- We model spread of true news in social network through developing an interventional model
- We model news diffusion as Multivariate Hawkes Process
- Interventions are learnt through policy optimization

Multivariate Hawkes Process

- It was developed by a statistic Professor named Alan Hawkes
- He introduced a family of probability models for the occurrence of series of events
- Provide probability of occurrence of an event by incorporating the entire history
- More accurate information will be fed into further account of things such as capital allocation, cash level prediction
- Applied to Social Networks, Insurance, Crime Prediction

What is Social Reinforcement?

- Refers to the process where acceptance and praise from others reinforces behaviors/preferences from others
- The authors propose a model to gain feedback behaviors of individual to make better incentivization policies
- Rewards in social media (likes, shares) are form of acceptance and appreciations

Approach

- We estimate the response a user will get from social network upon sharing a post
- This indicates impact on diffusion which we can use for efficient allocation of incentive
- A user response depends on political bias and peer influence
- We model this as a MHP
- Our evaluation is on semi-synthetic and real world data
- The result demonstrates that our model outperforms alternatives

News Diffusion

- There’s some recent work on blocking posts of fake news but it violates user’s rights
- In Twitter, a user tweets more about certain types of news articles in the past, he is more likely to tweet more about similar articles in future
- His followers are more likely re-tweet similar information

News Diffusion

- Given a model of news diffusion, the aim is to increase spread of true news among people exposed to fake news
- The assumption is that increased exposure to true news will increase suspicion and mistrust for fake news leading to decrease in fake news spread in future
- Hence, the more a user is exposed to true news, the more she will tend to believe such news
- [Silverman, 2016] observed that 46% of fake news circulated on Facebook

Intervention Strategy

- We model user response as Multivariate Hawkes Process (MHP)
- In MHP, base is proportional to their political bias and interleave it with news diffusion process
- We estimate user’s initial political bias using community detection algorithm and propose a model to update the bias over time
- The goal is to learn intervention strategy by selecting optimal set of people to be incentivized

Policy Optimization

- We post the problem as policy optimization in Reinforcement Learning
- The Reward function is maximizing true news spread among people exposed to fake news
- We integrate MHP in RL framework

Policy Optimization

- Our setting is a cooperative multi agent RL problem
- In this, number of agents is large and state, action spaces are continues
- To evaluate performance of the model, we use two real world twitter datasets
- The author uses semi-synthetic data
- The results show adding intervention, increases spread of true news is beneficial for mitigating the impact of fake news

Related Work

- Most of related work focuses on detecting fake news using different features such as linguistic, demographic, community based
- They use network properties such as clustering coefficient, closeness between centrality, neighbor based features like number of followers, followee
- Stochastic point process have gained popularity in modeling user activities in social networks
- Hawkes process models have been used to capture self exciting nature of events

Related Work

- The author's approach to social reinforcement learning is related to previous work on multi-agent RL
- However, it was limited to small number of agents <50
- This work extends a large number of agents >1000

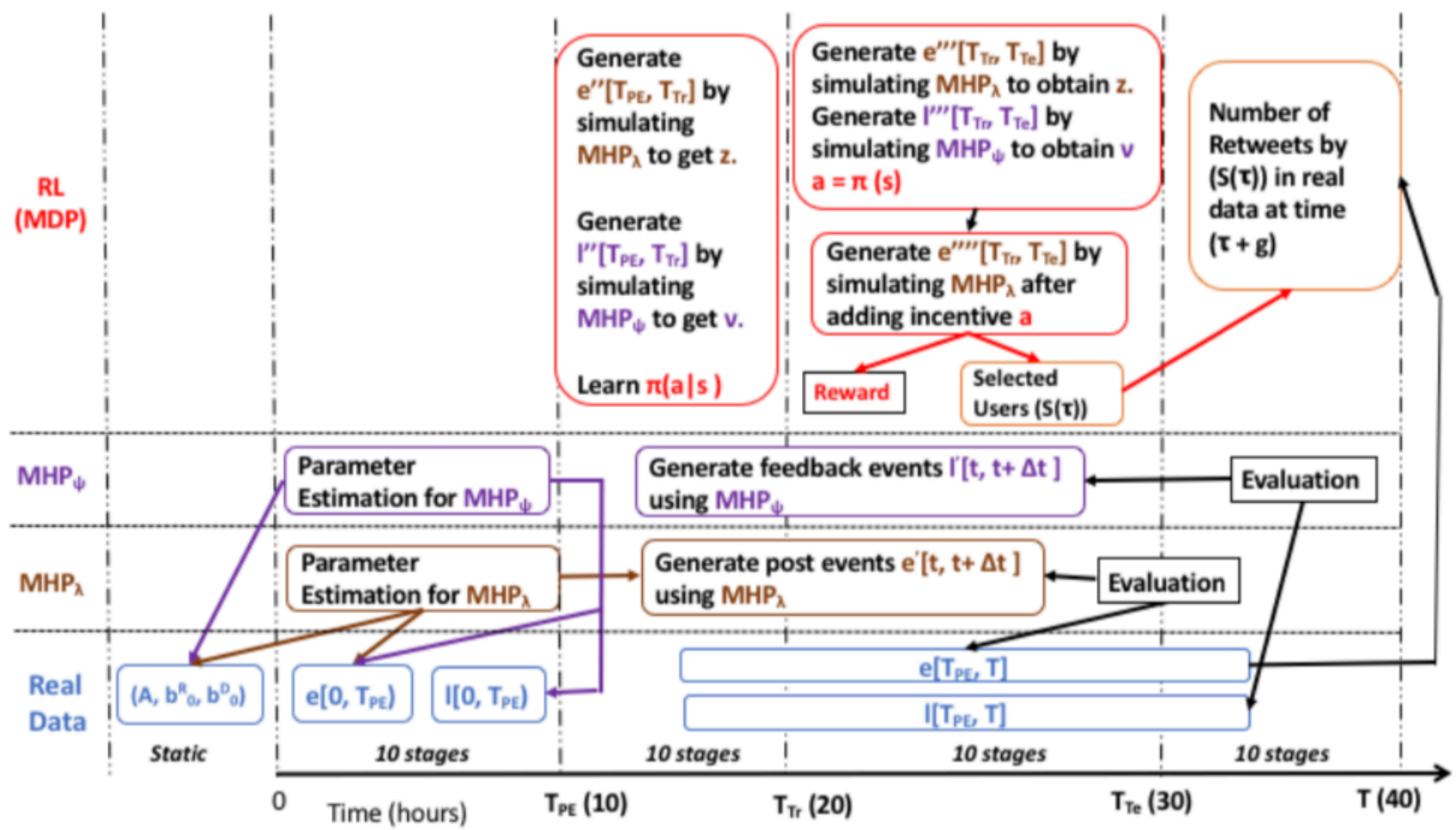
Evaluation

- To compare the performance of different methods, we consider the reward along with the fraction of users exposed to fake news that become ex- posed to true news

Reward

- We use the correlation between exposures to fake and true news to quantify our objective that people exposed more to fake news are also exposed more to true news.

Complete Framework



Experiments

- We evaluate a policy on test dataset
- We simulate the network under that policy 10 times, and report the average
- Vanilla MHP - Policy is function of user events (tweets)
- Exposure-based Policy EXP - To mitigate users who shared more posts related to fake news in past

Results

- There’s correlation between fake and true news exposure that is higher for MHP-U, it maximizes distinct number of users exposed to fake news
- There’s change in performance with respect to ratio of decay parameter for true and fake news diffusion
- When this ration increases, the performance decreases exponentially

Results

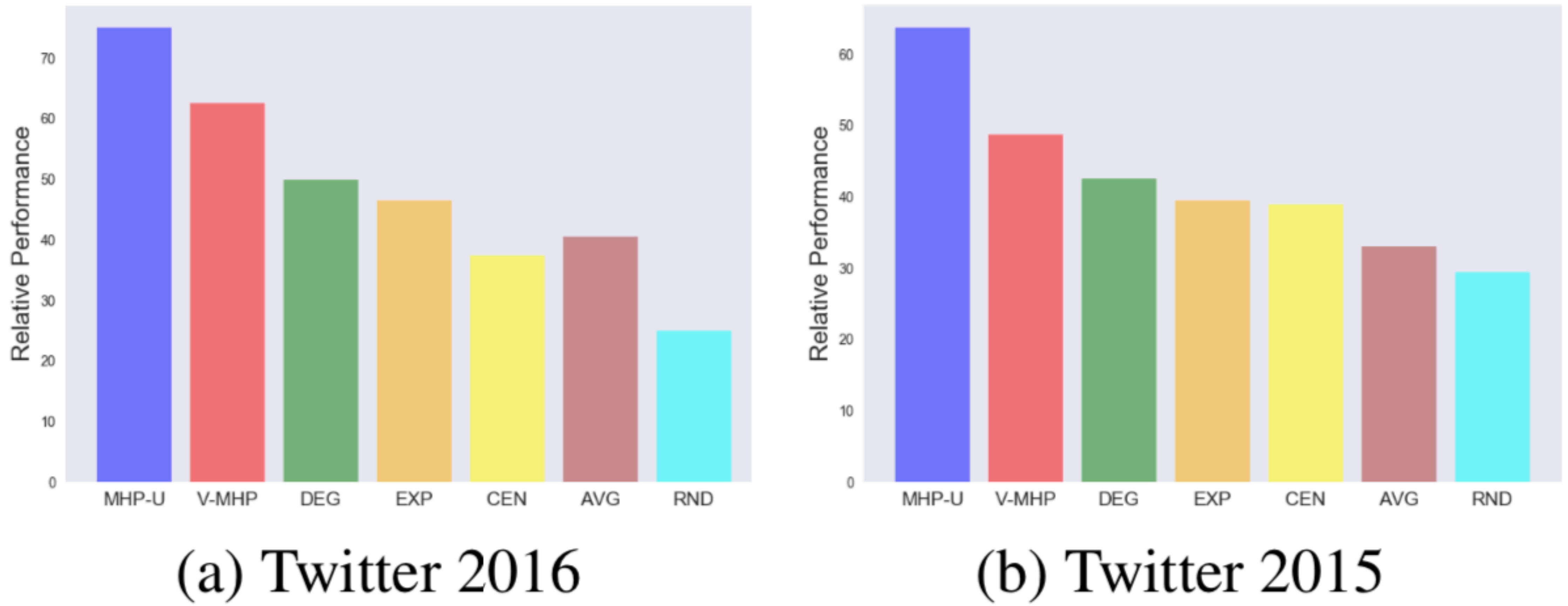


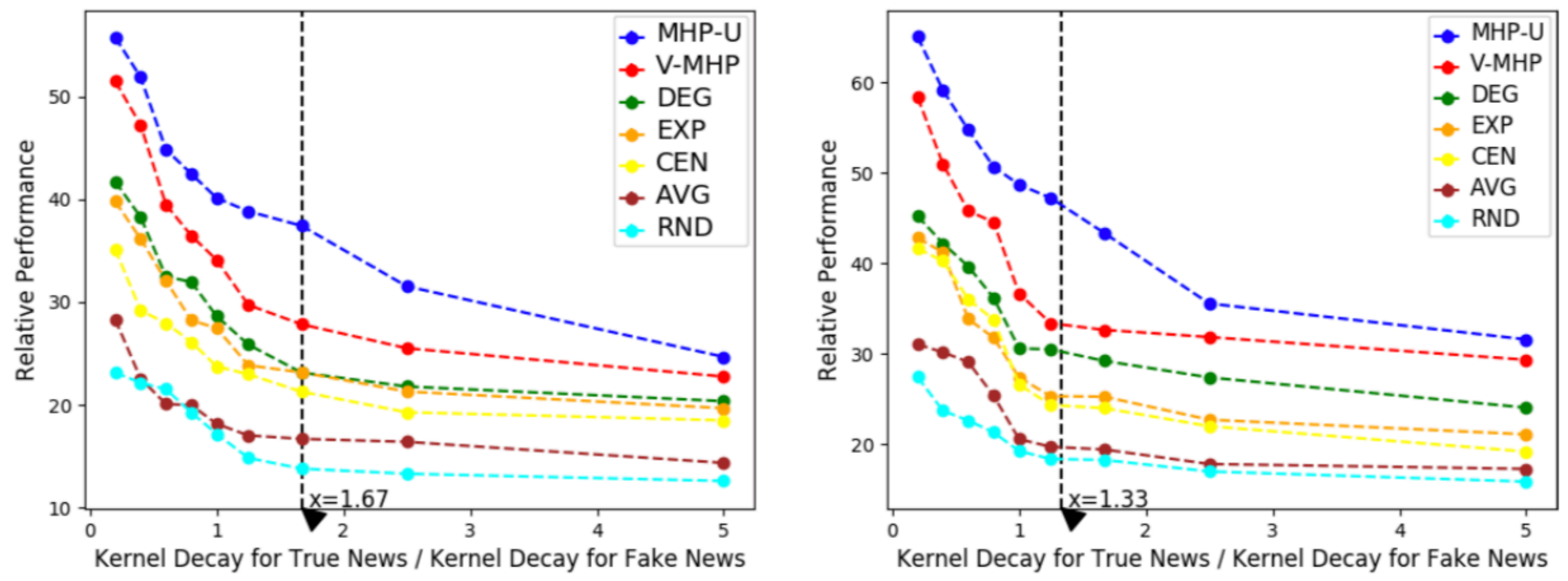
Figure 3: Relative Performance on Twitter Datasets

The twitter dataset shows the relative performance of different methods.

The correlation between fake and true news exposures is higher for MHP-U, and it also maximizes distinct number of users exposed to fake news.

It provides incentives to helps the increase of spread in true news even among the people exposed to fake news.

Results



(a) Twitter 2016

(b) Twitter 2015

Figure 4: Relative Performance vs Ratio of Decay

The Results shows a change in performance with respect to the ratio of decay parameter for true and fake news diffusion.

Summary

-The paper presents a social reinforcement learning approach

-It can be used to combat dissemination of fake news by learning how to incentivize users to spread more true news

-The key insight is: estimate likely feedback for each user based on both their network structure and the political bias of their followers

- Next, they combine the estimates with the observed events while learning an incentivization policy

-Experiments show that the above approach achieves better performance in terms of expected reward and number of distinct mitigated users