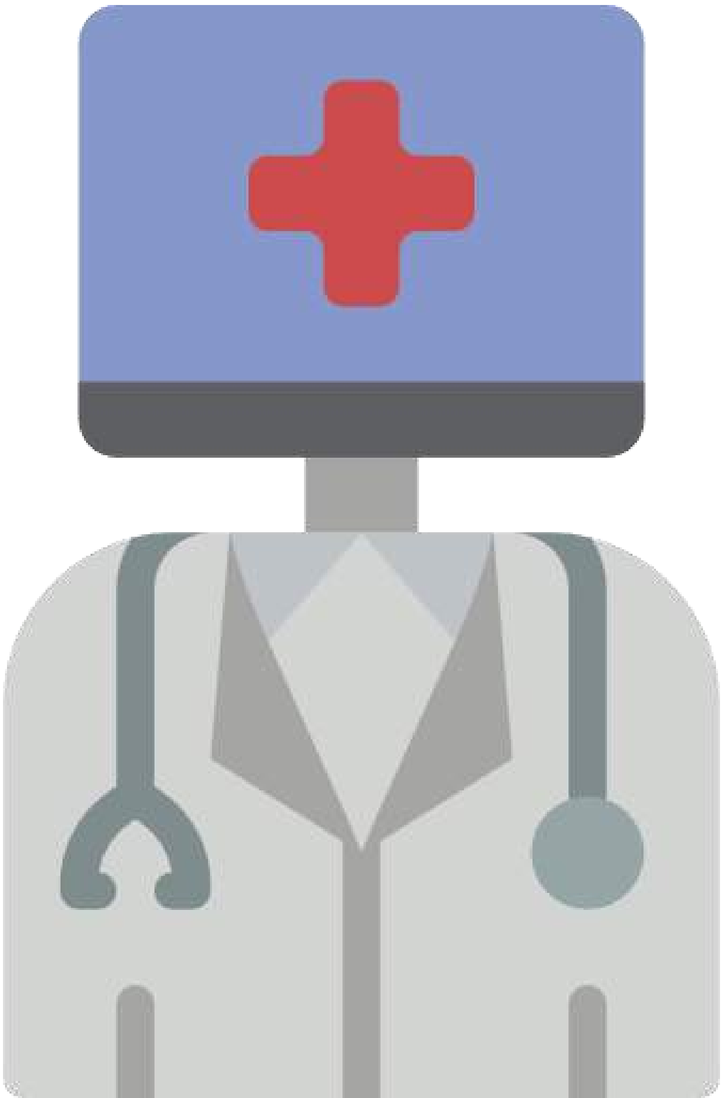


Reinforcement Learning
for Optimal Sepsis Treatment

Youngsuk Lee | Wonjin Woo | Seokhyeon Lee

CREDITS

Nikil Dutt
University of California, Irvine

Amir M. Rahmani
University of California, Irvine

Rui Cao
University of California, Irvine

Yong Huang
University of California, Irvine

Youngsuk Lee
Hanyang Univ.

Wonjin Woo
Hanyang Univ.

Seokhyeon Lee
Dankook Univ.

CONTENTS

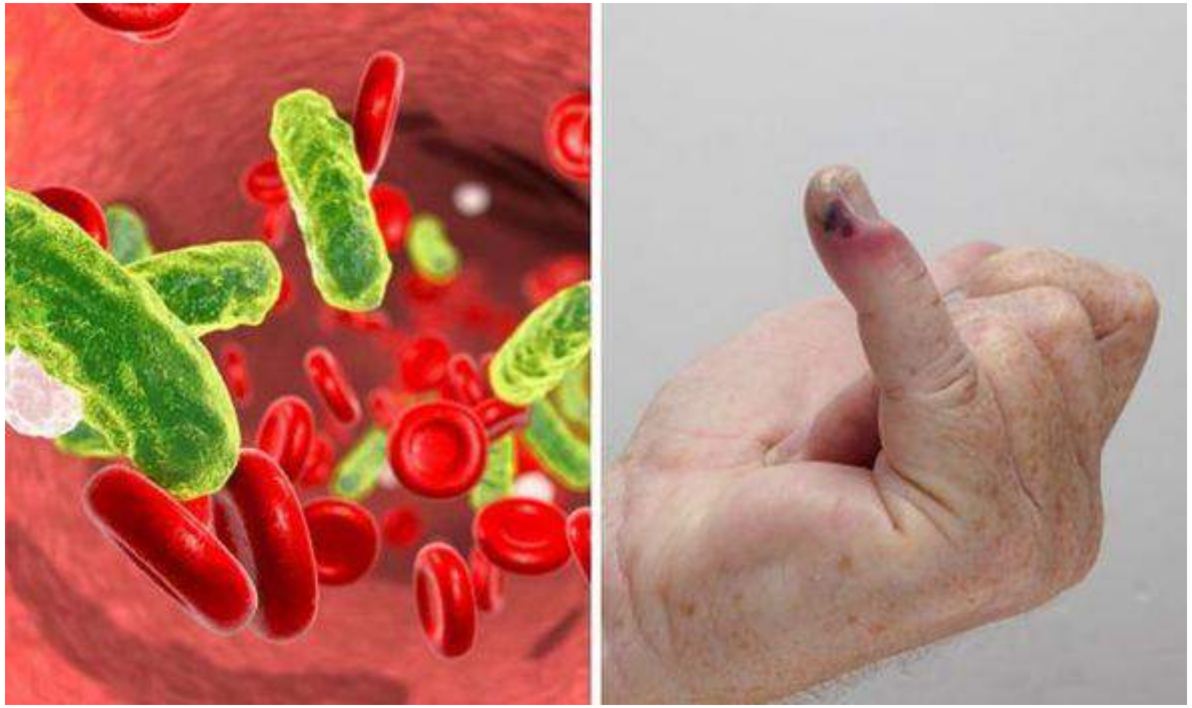
01 Sepsis Treatment

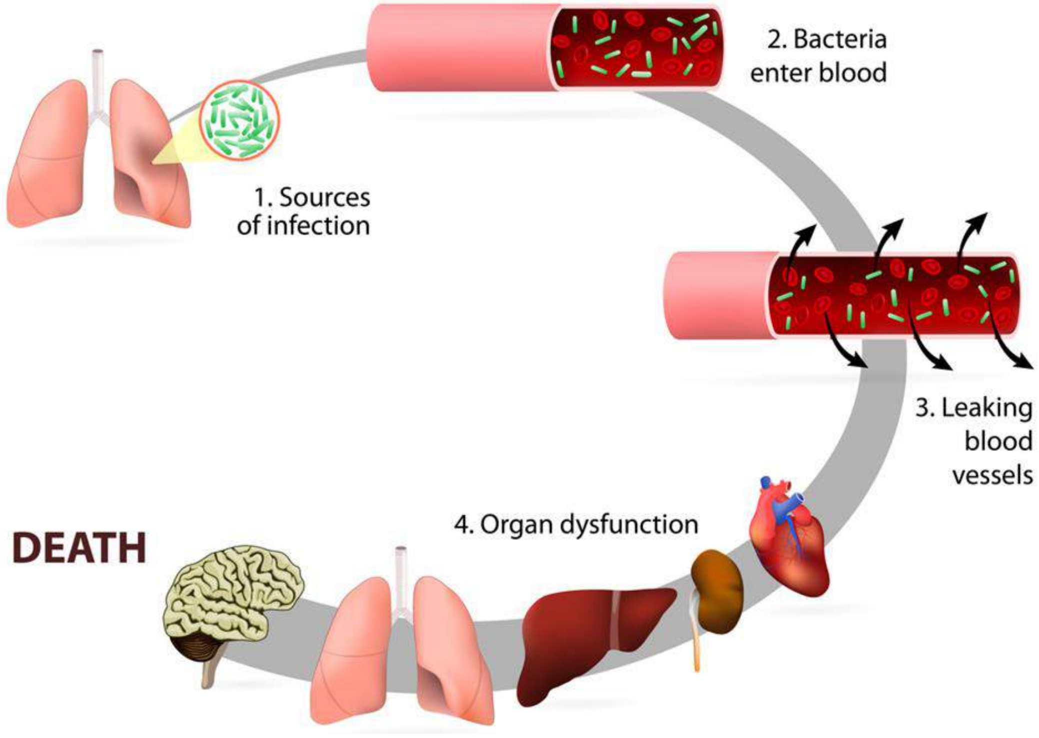
02 About Our Project

03 Achievements

04 Future Work & Expectation

Sepsis

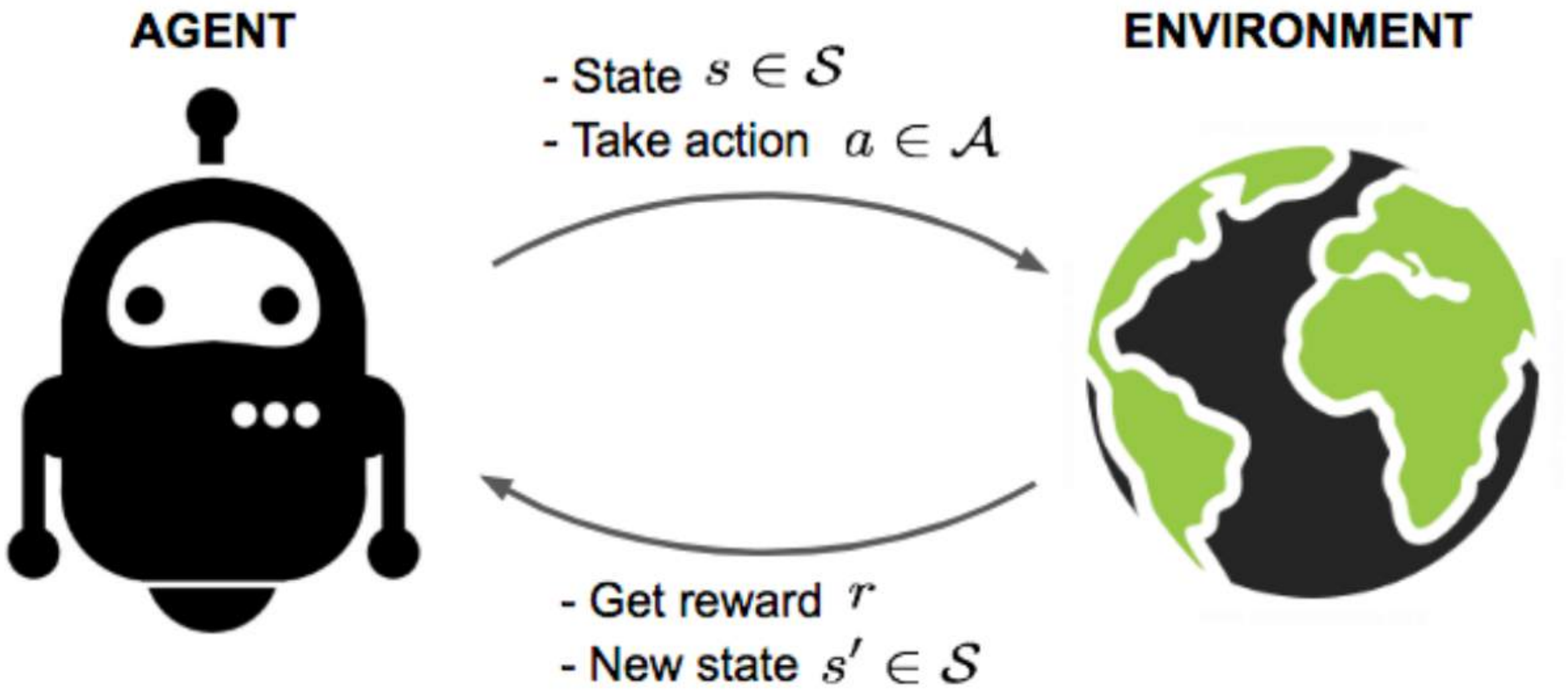



2. Bacteria enter blood
1. Sources of infection
3. Leaking blood vessels
4. Organ dysfunction
DEATH

- How to treat?

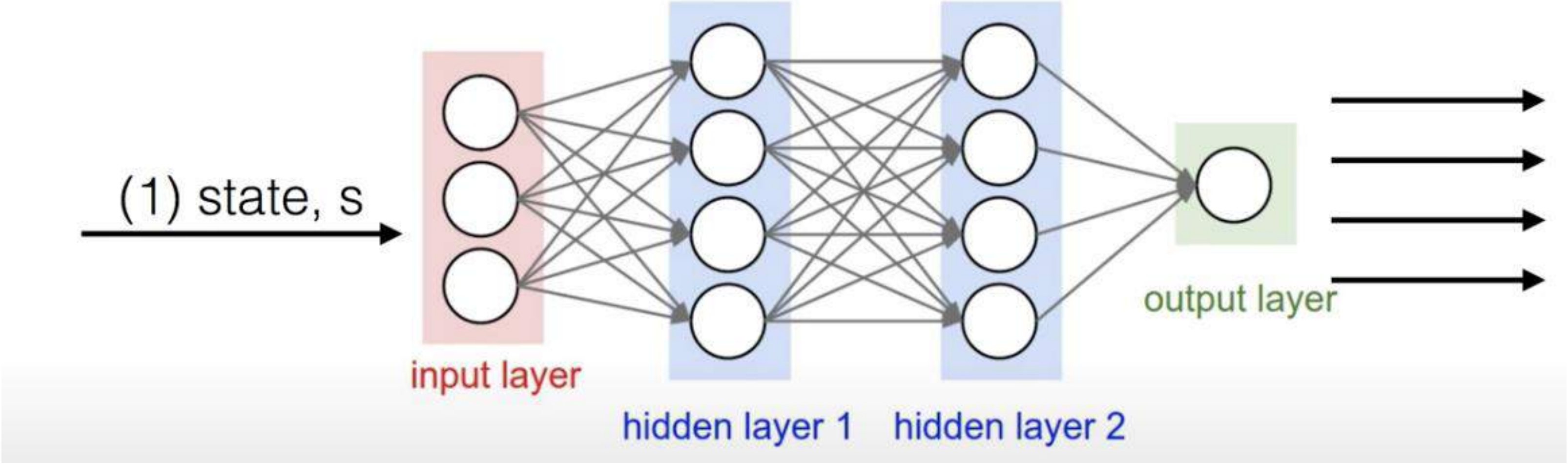
“Reinforcement Learning”

Reinforcement Learning



Q-learning

- Model-free & off-policy
- Q-value: $Q_\pi = E_\pi[R_{t+1} + \gamma R_{t+2} + \cdots + \mid S_t = s, A_t = \mathrm{a}]$



Q-learning

- Model-free & off-policy
- Q-value: $Q_\pi = E_\pi[R_{t+1} + \gamma R_{t+2} + \cdots + \mid S_t = s, A_t = \mathrm{a}]$

Instability & divergence problem

- DQN algorithm in Reinforcement Learning

Q-learning

- Model-free & off-policy

- Q-value:

$$Q_\pi = E_\pi[R_{t+1} + \gamma R_{t+2} + \cdots + \mid S_t = s, A_t = \mathrm{a}]$$

Instability & divergence problem

solve

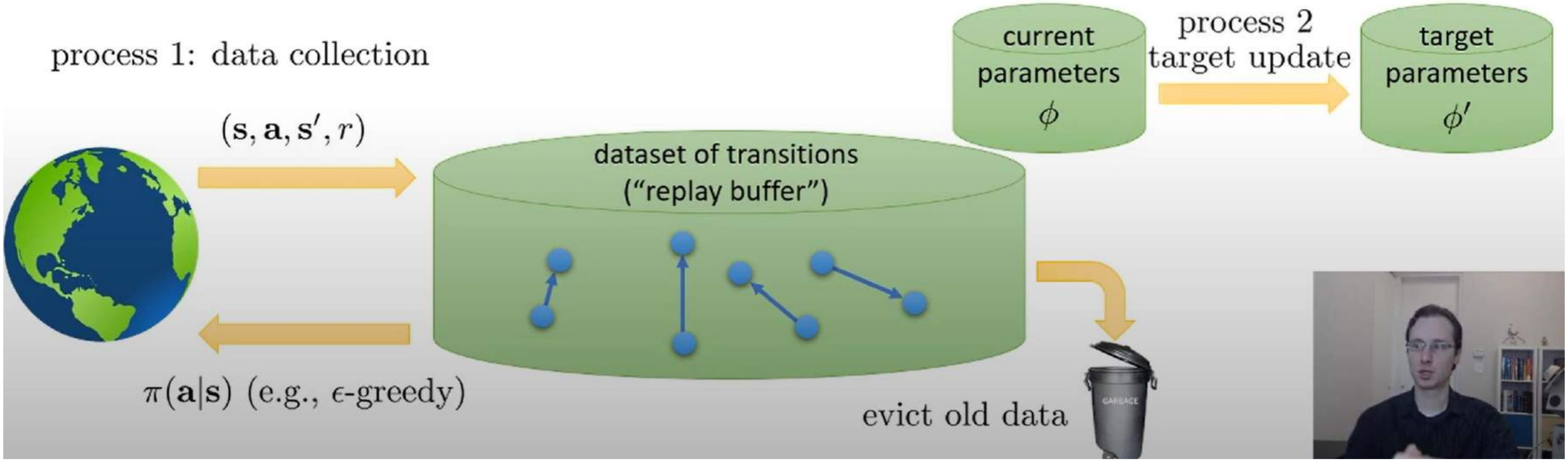
DQN

- Experience replay buffer

- Periodically updated Q-value

- DQN algorithm in Reinforcement Learning

DQN



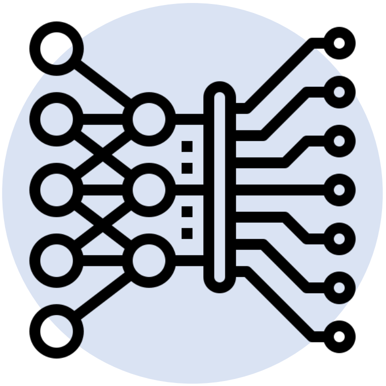
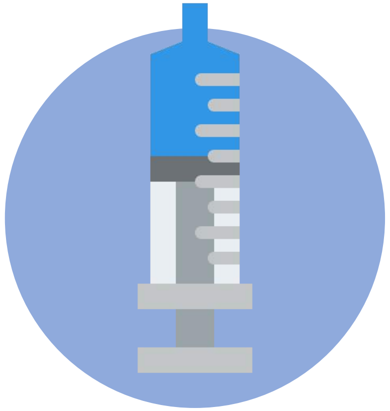
"

AI clinician: Reinforcement Learning for Optimal Sepsis Treatment

"

02 About Our Project

- Our Approach

Model	Reward	Goal
		
DQN	SOFA Score	Find the best Dosage Policies

Data Preprocessing

List of databases

Name	Owner	Encoding	Collate	Ctype	Access privileges	Size
mimic	postgres	UTF8	C	C		49 GB
postgres	zerostone	UTF8	C	C		7997 kB

Medical Information Mart for Intensive Care

03 Achievements

- Preprocessing

Data Preprocessing

- abx.csv
- ce010000.csv
- ce1000020000.csv
- ce2000030000.csv
- ce3000040000.csv
- ce4000050000.csv
- ce5000060000.csv
- ce6000070000.csv
- ce7000080000.csv
- ce8000090000.csv
- ce90000100000.csv
- culture.csv
- demog.csv
- fluid_cv.csv
- fluid_mv.csv
- labs_ce.csv
- labs_le.csv
- mechvent.csv
- microbio.csv
- preadm_fluid.csv
- preadm_uo.csv
- uo.csv
- vaso_cv.csv
- vaso_mv.csv

① Relevant Variables

- sepsis_final_data_RAW_withTimes.csv
- sepsis_final_data_withTimes.csv

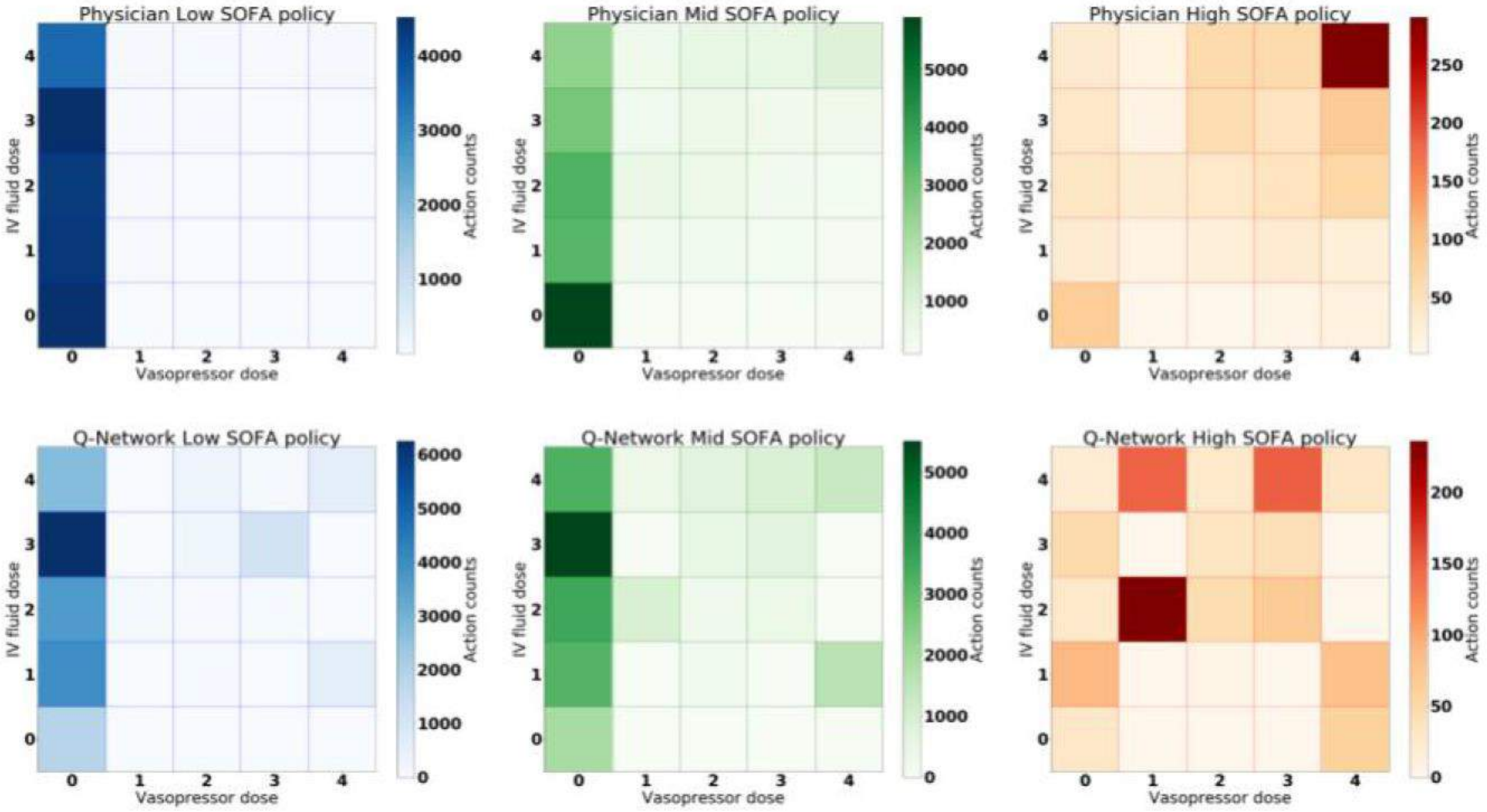
② MDP

- acuity_scores.csv
- test_mortality_tuple
- test_set_tuples
- train_set_tuples
- val_set_tuples

Will be used as
Train / Valid Set
in DQN

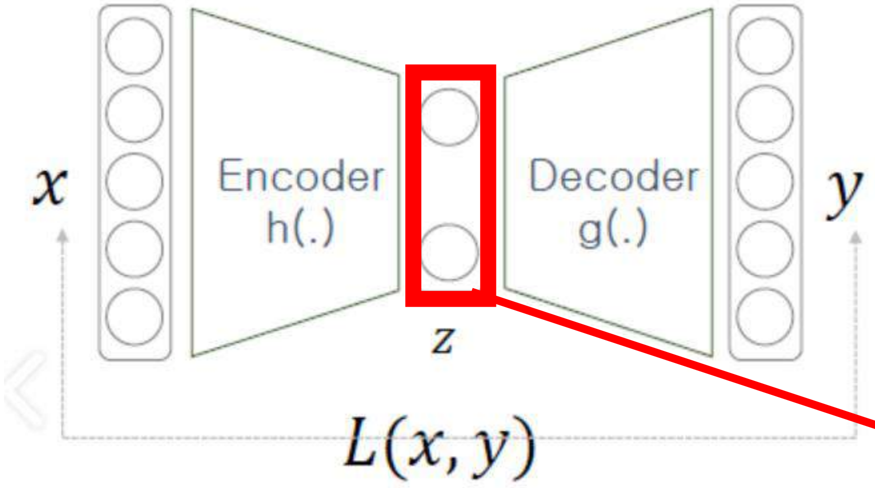
03 Achievements

- Actions

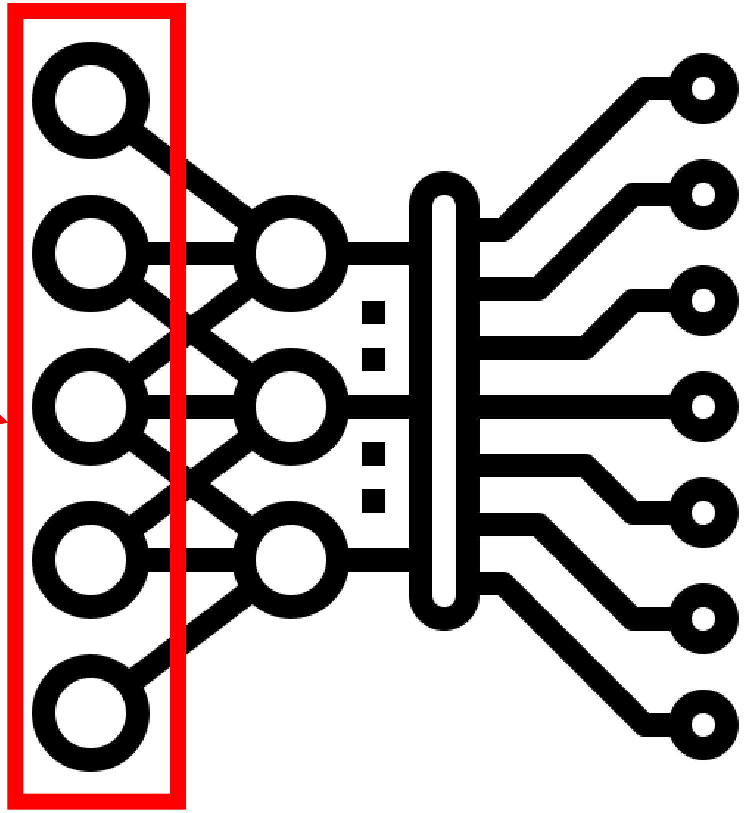


03 Achievements

- Auto-Encoding



Auto-Encoding



DQN

03 Achievements

- Model Construction

DQN

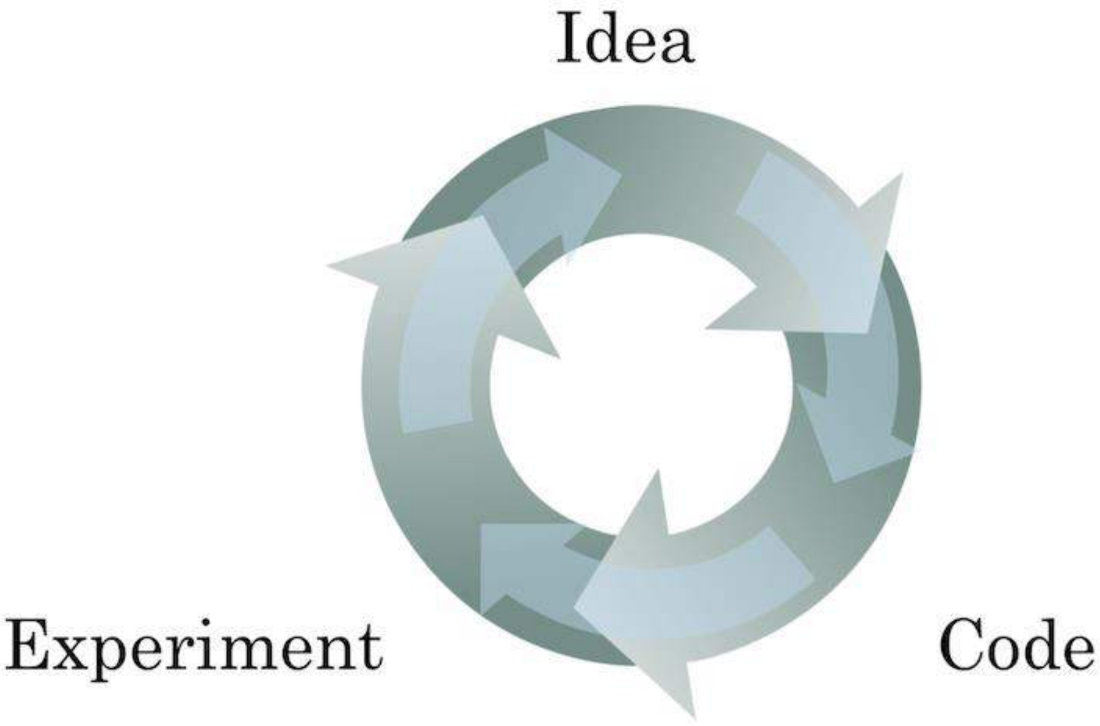
```
def train(self, replay_buffer):
    # Sample replay buffer
    state, action, next_state, reward, done, obs_state, next_obs_state = replay_buffer.sample()

    # Compute the target Q value
    with torch.no_grad():
        q_curr = self.Q(next_state)

        # Use large negative number to mask actions from argmax
        next_action = q_curr.argmax(1, keepdim=True)

        q_target = self.Q_target(next_state)
        target_Q = 10*reward + done * self.discount * q_target.gather(1, next_action).reshape(-1, 1)

    # Get current Q estimate
    current_Q = self.Q(state)
    current_Q = current_Q.gather(1, action)

    # Compute Q loss
    Q_loss = F.smooth_l1_loss(current_Q, target_Q)

    # Optimize the Q
    self.Q_optimizer.zero_grad()
    Q_loss.backward()
    self.Q_optimizer.step()

    # Update target network by polyak or full copy every X iterations.
    self.iterations += 1
    self.maybe_update_target()
    return reward, Q_loss.item()
```

- Moduling Evaluation Function

importance sampling

$$
\begin{aligned}
E_{x\sim p(x)}[f(x)] &= \int p(x)f(x)dx \\
&= \int \frac{q(x)}{q(x)} p(x)f(x)dx \\
&= \int q(x)\frac{p(x)}{q(x)} f(x)dx \\
&= E_{x\sim q(x)}\left[\frac{p(x)}{q(x)} f(x)\right]
\end{aligned}
$$

- A lot of Cycles



Continuing Experiment

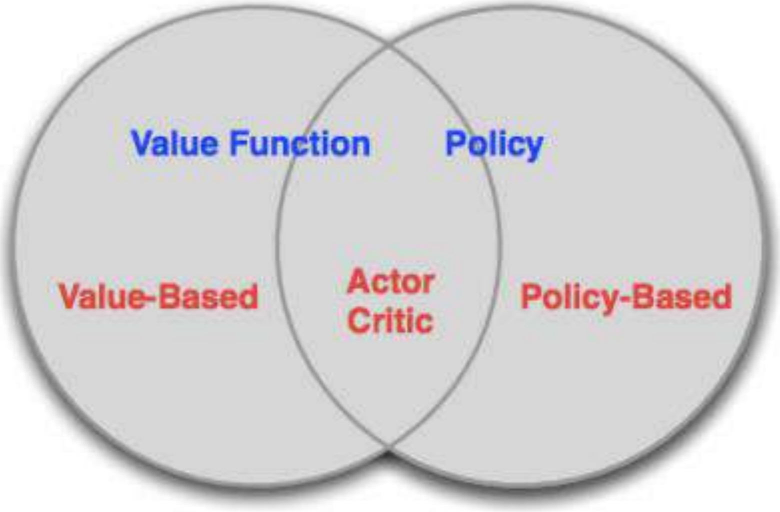
&&

Hyper Parameter Tuning

- Create Models Using other algorithms

Policy Gradient Algorithm

$$\nabla_\theta J(\theta) = \nabla_\theta \sum_{s\in S} d^\pi(s) \sum_{a\in A} Q^\pi(s,a)\pi_\theta(a|s)$$



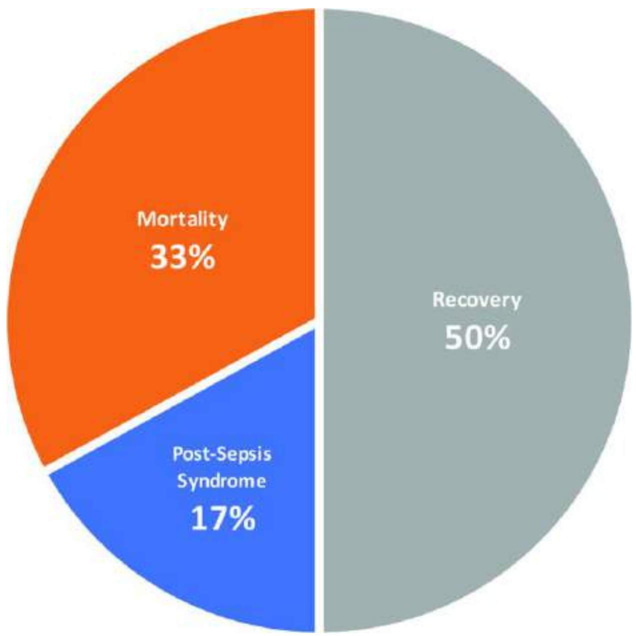
Policy-Based Algorithms : Reinforcement …

Value-Based Algorithms : DQN …

Policy-Based + Value-Based Algorithm : Actor Critic …

04 Future Work & Expectation

Sepsis is diagnosed in at least 1.7 million adults annually in the United States, according to the Centers for Disease Control and Prevention. About 270,000 Americans die from sepsis every year, and 1 in 3 patients who die in hospitals are diagnosed with sepsis, the CDC says.



Reducing small rates of mortality
can **save a lot of people**
who suffer from sepsis
(About 8,200 Americans per 1 percent)


THANK YOU!

Q&A

Paper Resources

Deep Reinforcement Learning for Sepsis Treatment (2017), Aniruddh Raghu

Reinforcement Learning for Sepsis Treatment: Baselines and Analysis (2019), Aniruddh Raghu

Improving Sepsis Treatment Strategies by Combining Deep and Kernel-Based Reinforcement Learning (2019), Xuefeng Peng

Image Resources

Stanford University

Health Headers

Molmed

EXPRESS

EMS1

A (Long) Peek into Reinforcement Learning

Sung Kim Youtube channel

Other resources

A (Long) Peek into Reinforcement Learning

CDC

Stanford University

Health Headers

Molmed

Image Resources Link

https://www.express.co.uk/life-style/health/1369021/sepsis-symptoms-what-causes-sepsis-how-to-prevent-sepsis-evg

https://www.ems1.com/ems-products/medical-monitoring/articles/what-ems-needs-to-know-about-new-pediatric-sepsis-guidelines-u0gCT72EzloEI8Zd/

https://lilianweng.github.io/lil-log/2018/02/19/a-long-peek-into-reinforcement-learning.html

http://hunkim.github.io/ml/

https://web.stanford.edu/class/cs234

https://www.healthleadersmedia.com/clinical-care/new-data-sepsis-prevalence-and-costs-astonished-dhhs-researchers
https://molmed.biomedcentral.com/articles/10.1186/s10020-019-0132-z

https://deepinsight.tistory.com/126