Mahidol University
Faculty of Medicine Ramathibodi Hospital
Department of Clinical Epidemiology and Biostatistics

Reinforcement learning in health care

RADI608: Data Mining and Machine Learning
RADI602: Data Mining and Knowledge Discovery

Lect. Anuchate Pattanateepapon, D.Eng.

Section of Data Science for Healthcare

Department of Clinical Epidemiology and Biostatistics

Faculty of Medicine Ramathibodi Hospital, Mahidol University

© 2022

Wisdom of the Land

Examples of the reinforcement learning in health care

- Deep Reinforcement Learning to Detect Brain Lesions on MRI (Joseph N Stember, 2020)
- Improving Sepsis Treatment Strategies by Combining Deep and Kernel-Based Reinforcement Learning (Xuefeng Peng, 2019)
- Deep Reinforcement Learning for Surgical Gesture Segmentation and Classification (Daochang Liu, Tingting Jiang, 2018)
- Reinforcement Learning Strategies for Clinical Trials in Nonsmall Cell Lung Cancer. Biometrics (Y. Zhao, 2011)
- Pruning an ensemble of classifiers via reinforcement learning (Ioannis Partalas, Grigorios Tsoumakas and Ioannis Vlahavas, 2009)

Deep Reinforcement Learning to Detect Brain Lesions on MRI: a proof-of-concept application of reinforcement learning to medical images

Cornell University

arXiv.org > cs > arXiv:2008.02708

Search...

Help | Adv

Computer Science > Artificial Intelligence

[Submitted on 6 Aug 2020]

Deep reinforcement learning to detect brain lesions on MRI: a proof-of-concept application of reinforcement learning to medical images

Joseph Stember, Hrithwik Shalu

Deep Reinforcement Learning to Detect Brain Lesions on MRI

- AI in radiology is hindered chiefly by:

  1) Requiring large annotated data sets.

  2) Non-generalizability that limits deployment to new scanners /institutions.

  3) Inadequate explain ability and interpretability.

- They believe that reinforcement learning can address all three shortcomings, with robust and intuitive algorithms trainable on small datasets.

Deep Reinforcement Learning to Detect Brain Lesions on MRI

Methodology

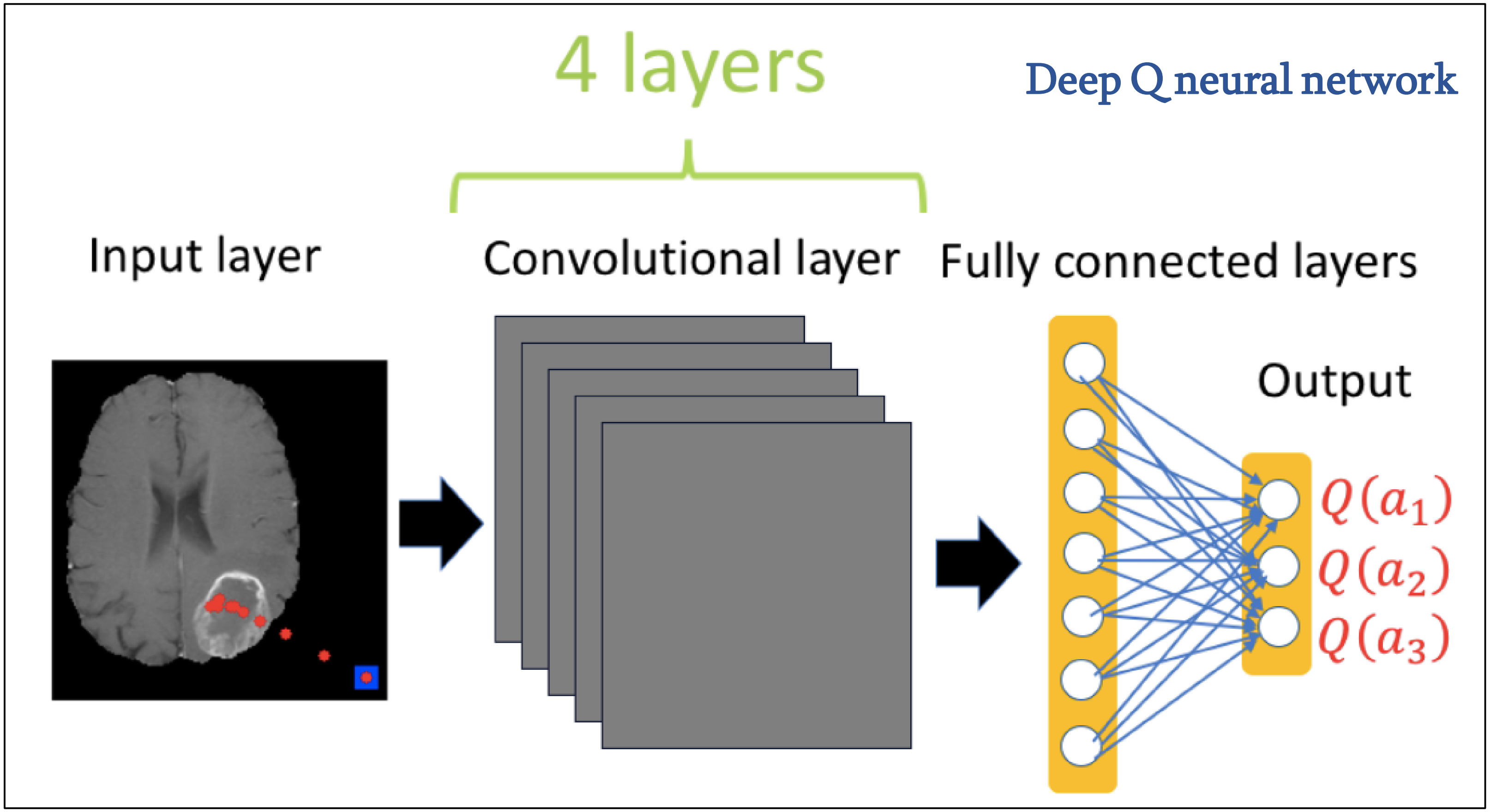
- **Objective :**
  - To predict tumor location in MRI by using reinforcement learning
  - To compare predicting accuracy of reinforcement learning with supervised deep learning method.
- **Method :**

  The images they used for training and testing came from the BraTS high grade glioma database. From that database's T1-weighted post-contrast 3D image volumes, they randomly selected 100 2D image slices. They employed a 70/30 training-testing split, using the first 70 of these images for training. They trained for 300 episodes.

Deep Reinforcement Learning to Detect Brain Lesions on MRI

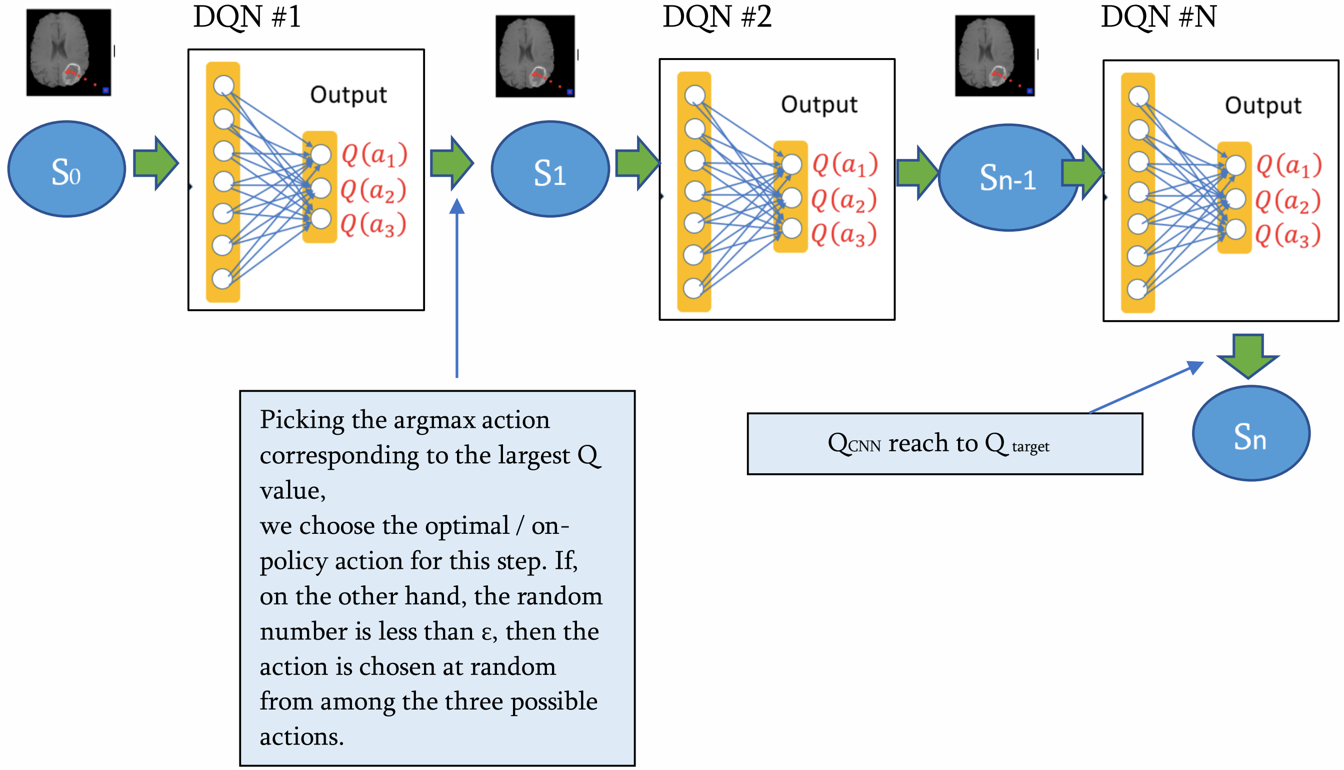
Environment	The 2D slice of a post-contrast T1-weighted brain MRI containing a glioblastoma multiforme (GBM) lesion with overlaid gaze plot (eye tracking from radiologist) for that image.
Action	Point moving between pixels in the image, more specifically between points on the gaze plot.
State	The pixel that the point sit, related to the lesion from imaging
Reward	Integer value from the pixel position and direction of the point related to the tumor Example +2 : if agent is within lesion and staying still. —1.5 : if agent is outside lesion and moving backward
Policy	The agent to move toward the tumor as quickly as possible and stay there to mark / predict that lesion.

Deep Reinforcement Learning to Detect Brain Lesions on MRI



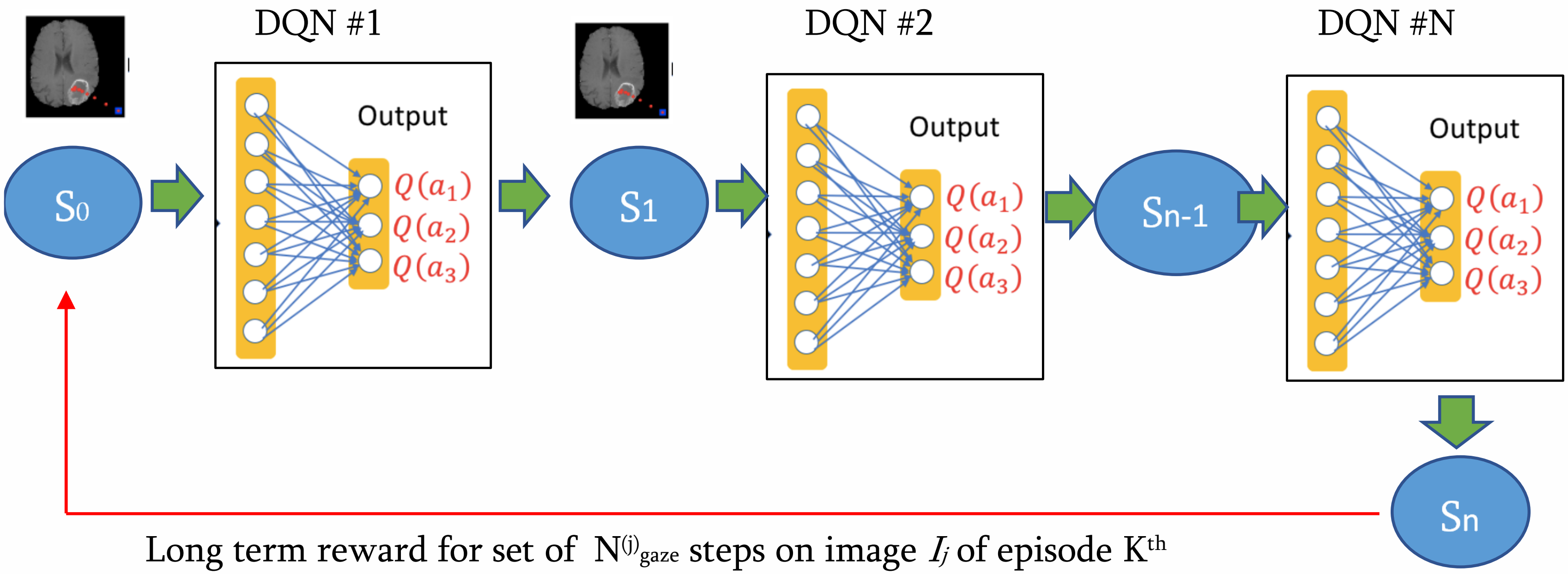
Deep Reinforcement Learning to Detect Brain Lesions on MRI

First episode



Deep Reinforcement Learning to Detect Brain Lesions on MRI

Second to Kth episode



Long term reward for set of $N^{(j)}_{gaze}$ steps on image I_j of episode Kth
(In this article run 300 episodes)

Improving Sepsis Treatment Strategies by Combining Deep and Kernel-Based Reinforcement Learning

Improving Sepsis Treatment Strategies by Combining Deep and Kernel-Based Reinforcement Learning

Xuefeng Peng, MSc.[1], Yi Ding, MSc.[2], David Wihl, A.L.B.[1], Omer Gottesman, PhD[1], Matthieu Komorowski, MD[3], Li-wei H. Lehman, PhD[4], Andrew Ross, MSE[1], Aldo Faisal, PhD[3], Finale Doshi-Velez, PhD[1]
[1]Harvard University, Paulson School of Engineering and Applied Sciences, Cambridge, MA
[2]Harvard University, T.H. Chan School of Public Health, Cambridge, MA
[3] Imperial College London, London, UK
[4]MIT, Institute for Medical Engineering & Science, Cambridge, MA

Abstract

Sepsis is the leading cause of mortality in the ICU. It is challenging to manage because individual patients respond differently to treatment. Thus, tailoring treatment to the individual patient is essential for the best outcomes. In this paper, we take steps toward this goal by applying a mixture-of-experts framework to personalize sepsis treatment. The mixture model selectively alternates between neighbor-based (kernel) and deep reinforcement learning (DRL) experts depending on patient's current history. On a large retrospective cohort, this mixture-based approach outperforms physician, kernel only, and DRL-only experts.

Improving Sepsis Treatment Strategies by Combining Deep and Kernel-Based Reinforcement Learning

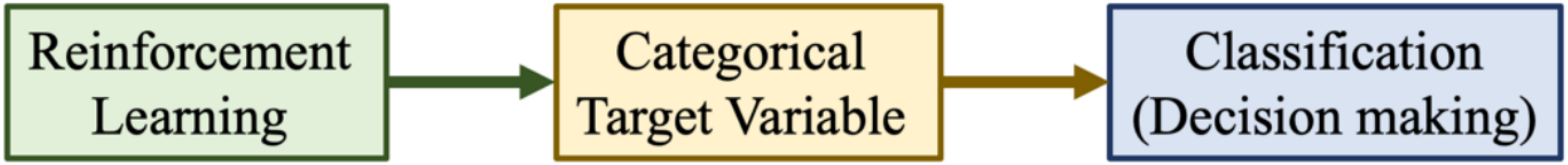
Background:

- Sepsis is emergency medical condition which require rapid treatment, but
- The challenge is each patient has individual variation responding to treatment.
- This paper focused on hypovolemia and vasodilation which are common symptoms among patients with sepsis and need intravenous (IV) fluid and vasopressors (VP) to cure these symptoms.

Objective:

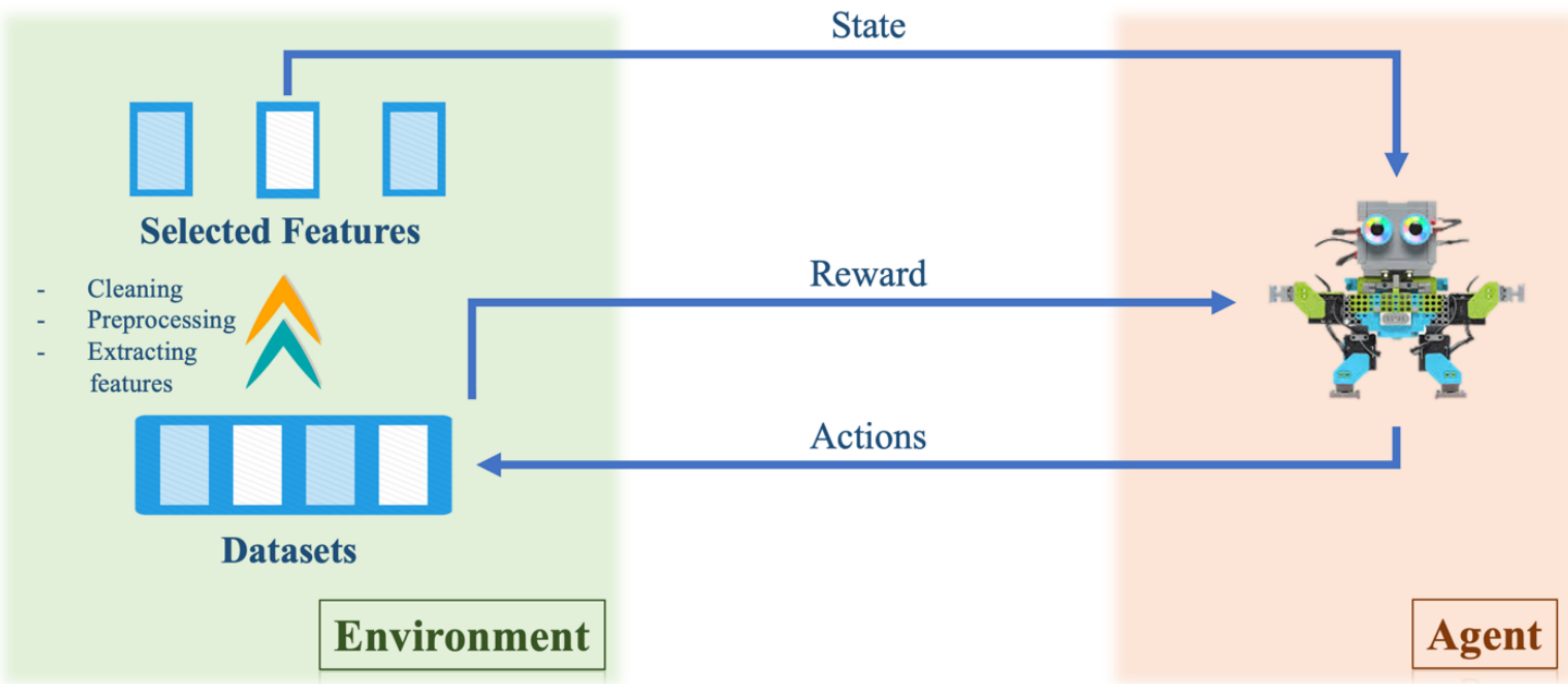
Find personalize sepsis treatment strategies which are intravenous (IV) fluid and vasopressors (VP) to sepsis patient by applying reinforcement learning (optimal treatment policy).

Improving Sepsis Treatment Strategies by Combining Deep and Kernel-Based Reinforcement Learning



Mahidol University
Faculty of Medicine Ramathibodi Hospital
Department of Clinical Epidemiology and Biostatistics

Improving Sepsis Treatment Strategies by Combining Deep and Kernel-Based Reinforcement Learning



Wisdom of the Land

Improving Sepsis Treatment Strategies by Combining Deep and Kernel-Based Reinforcement Learning

Dataset:

- Cohort observational sequences critical care datasets (15,415 patients).
- Datasets were partitioned only 4 hours treatment window and selecting features by applying sepsis-3 criteria, then will got 50 attributes (selected features)
- Patents with missing values were excluded.
- Features did standardization and log transformation.
- Dataset was split into 75% training and 25% test set.

Improving Sepsis Treatment Strategies by Combining Deep and Kernel-Based Reinforcement Learning

Actions:

Focusing on administration of both intravenous (IV) fluid and vasopressors (VP). In cohort dataset, using IV and VP in 4 hours treatment window, the dosages for each medication were categorized into 5 groups. Thus, we got 5x5 as 25 actions in action spaces.

State:

Patient's history which has 50 features per patient was cumulative history for each patient in dataset. The dataset was encoded into patient states by using LSTM autoencoder. The LSTM had a single layer of 128 hidden units for both encoder and decoder—resulting in a state that consisted of 128 real-valued vectors.

Improving Sepsis Treatment Strategies by Combining Deep and Kernel-Based Reinforcement Learning

Reward:

A feedback for each action performed by the agent by focusing on reducing mortality among patients with sepsis. The mortality was defined as probability of mortality given a patient's current observations by using a two-layer neural network, and defining the reward as the change in the negative mortality log-odds of mortality between the current observations and the next observations.

Policy learning:

Using Mixture-of-Experts (MoE). The MoE was selectively alternates between neighbor-based (kernel) and deep reinforcement learning (DRL) experts depending on patient's current history (Next slide)

Improving Sepsis Treatment Strategies by Combining Deep and Kernel-Based Reinforcement Learning

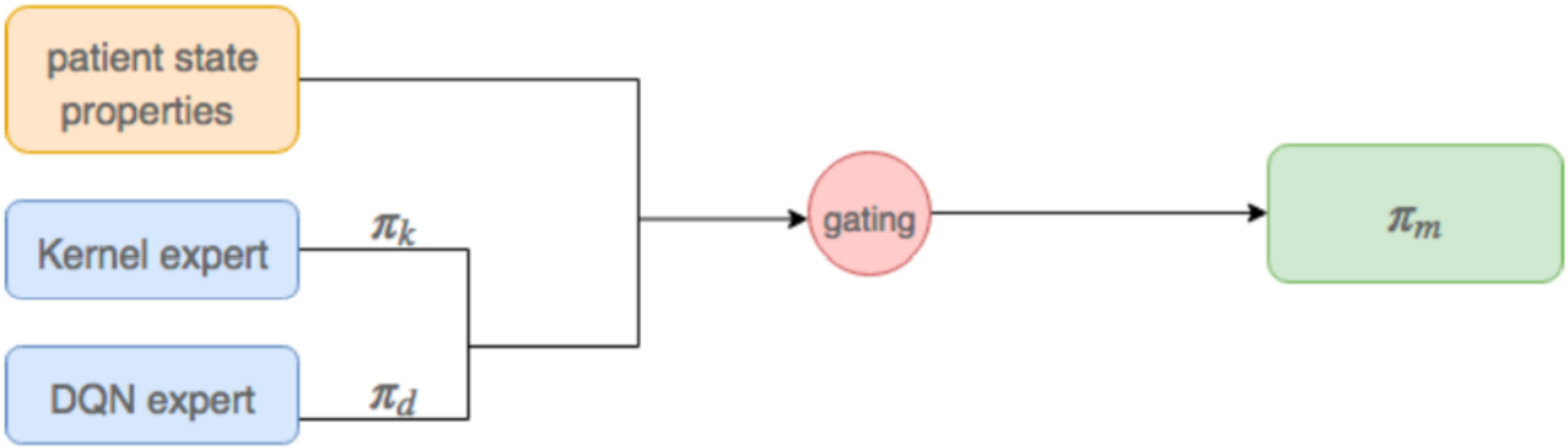


Fig. MoE, it produces a mixed policy via combining kernel

Improving Sepsis Treatment Strategies by Combining Deep and Kernel-Based Reinforcement Learning

Neighbor-based (kernel):

getting appropriate treatment by looking the nearest neighbors form the current states, identify the survivors (nearest neighbor), and choose actions that following by treatments performed on the nearby survivors.

Deep reinforcement learning (DRL):

applying to find a policy that outperforms the physician policy. The agent was trained to take action with the highest Q-value to achieve the goal of improving the survival.

Patient state properties:

finding the features that might be most useful for selecting between experts.

Deep Reinforcement Learning for Surgical Gesture Segmentation and Classification

Deep Reinforcement Learning for Surgical Gesture Segmentation and Classification

Daochang Liu and Tingting Jiang

National Engineering Lab for Video Technology, Cooperative Medianet Innovation Center, School of EECS, Peking University, Beijing 100871, China
{daochang, ttjiang}@pku.edu.cn

Abstract. Recognition of surgical gesture is crucial for surgical skill assessment and efficient surgery training. Prior works on this task are based on either variant graphical models such as HMMs and CRFs, or deep learning models such as Recurrent Neural Networks and Temporal Convolutional Networks. Most of the current approaches usually suffer from over-segmentation and therefore low segment-level edit scores. In contrast, we present an essentially different methodology by modeling the task as a sequential decision-making process. An intelligent agent

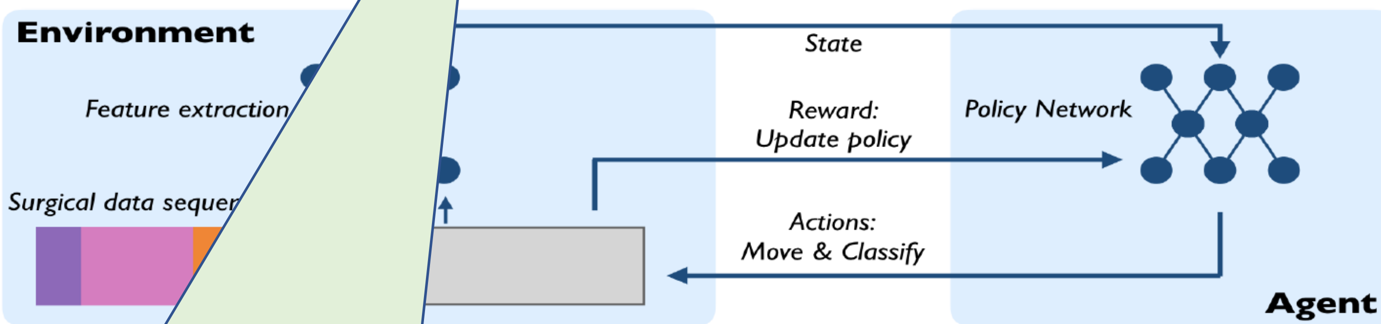
Deep Reinforcement Learning for Surgical Gesture Segmentation and Classification

Reinforcement Learning

Categorical Target Variable

Classification

- *model joint surgical gesture segmentation and classification as a sequential decision-making problem*
- *an agent perceives the environment, selects an action, and update its policy to maximize future rewards*



- For objective surgical skill assessment
- For improving efficiency and quality of surgery training

Mahidol University
Faculty of Medicine Ramathibodi Hospital
Department of Clinical Epidemiology and Biostatistics

Deep Reinforcement Learning for Surgical Gesture Segmentation and Classification

Reinforcement Learning

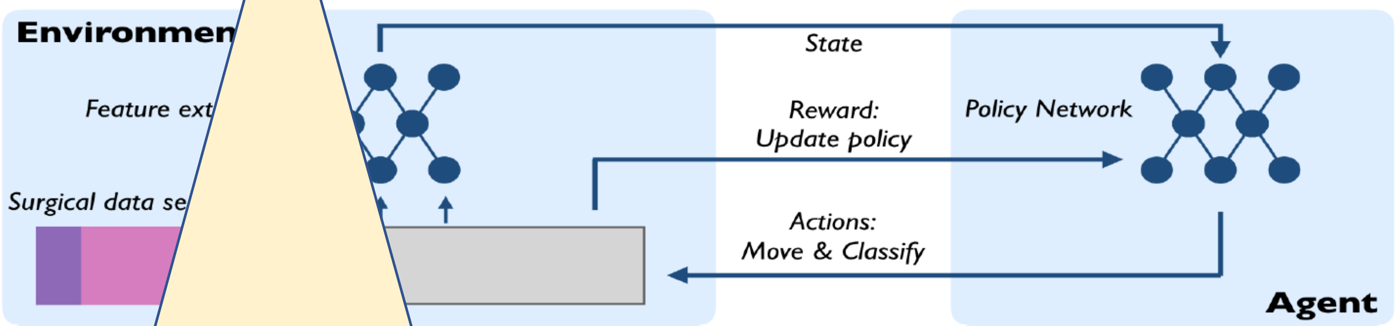
↓

Categorical Target Variable

↓

Classification

- *model joint surgical gesture segmentation and classification as a sequential decision-making problem*
- *an agent erceives the environment, selects an action, and update its policy to maximiz ture rewards*



- To segment robotic kinematic data
- To segment video sequence
- To classify segmented pieces into surgical gestures

Deep Reinforcement Learning for Surgical Gesture Segmentation and Classification

Reinforcement Learning

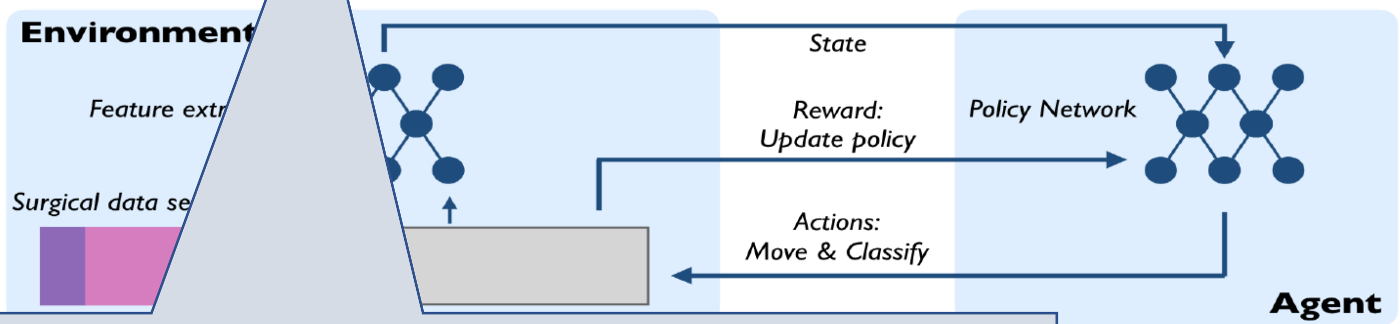
↓

Categorical Target Variable

↓

Classification

- *model joint surgical gesture segmentation and classification as a sequential decision-making problem*
- *an agent pe ceives the environment, selects an action, and update its policy to maximize re rewards*



Such as:

- Reaching for the needle
- Orienting needle
- Pushing needle through the tissue

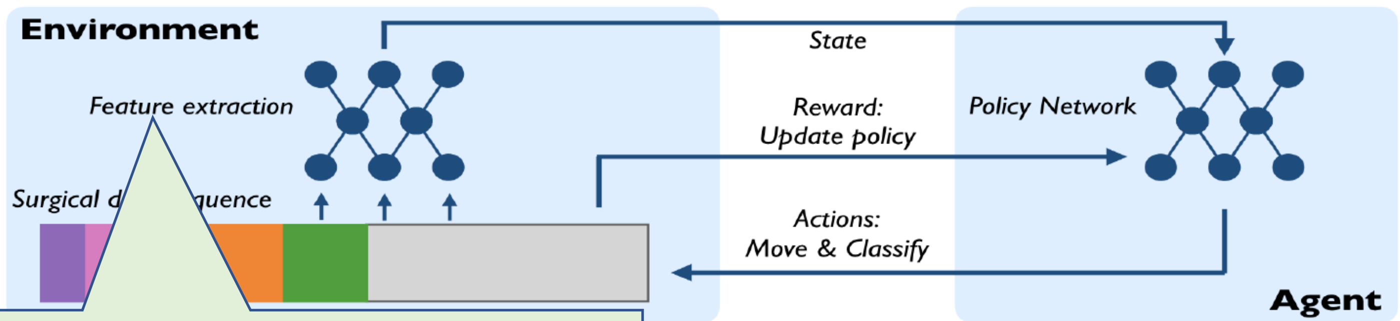
Deep Reinforcement Learning for Surgical Gesture Segmentation and Classification

Reinforcement Learning

Categorical Target Variable

Classification

- *model joint surgical gesture segmentation and classification as a sequential decision-making problem*
- *an agent perceives the environment, selects an action, and update its policy to maximize future rewards*



Visual features

- extract from surgery video frames or kinematic data frames

- collect from surgical robots

Deep Reinforcement Learning for Surgical Gesture Segmentation and Classification

Reinforcement Learning

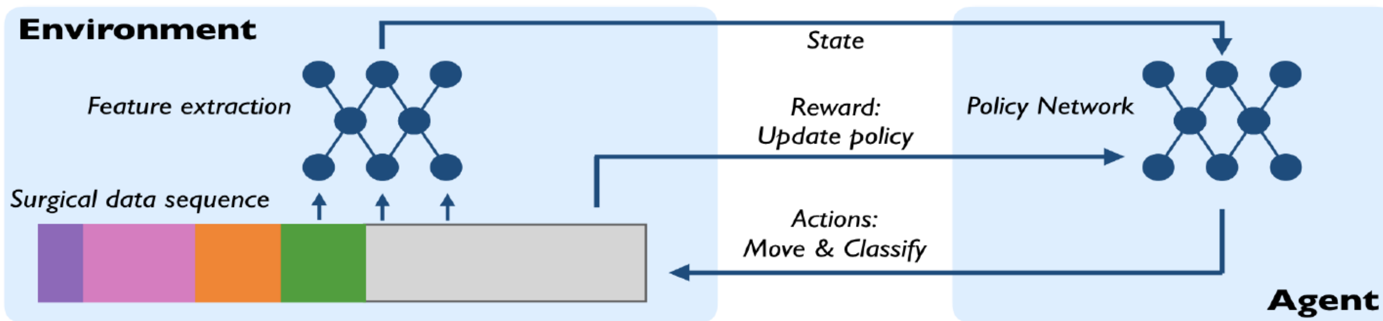
↓

Categorical Target Variable

↓

Classification

- *model joint surgical gesture segmentation and classification as a sequential decision-making problem*
- *an agent perceives the environment, selects an action, and update its policy to maximize future rewards*



Action: decide how far to move forward and choose which class label to give

State: combination of high-level representation and other auxiliary information, which assists the agent to make a better decision

Reward: a numerical feedback for each action performed by the agent

Policy learning: a multilayer perceptron to model the policy

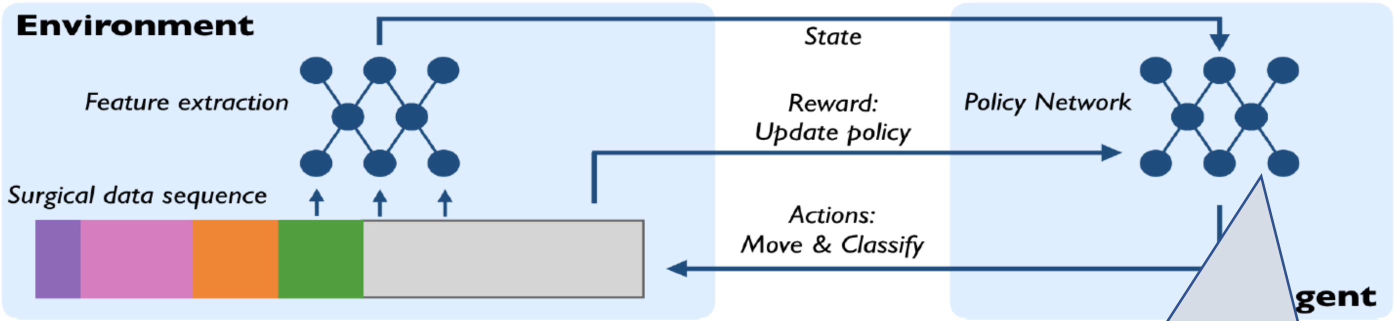
Deep Reinforcement Learning for Surgical Gesture Segmentation and Classification

Reinforcement Learning

Categorical Target Variable

Classification

- *model joint surgical gesture segmentation and classification as a sequential decision-making problem*
- *an agent perceives the environment, selects an action, and update its policy to maximize future rewards*



Action: decide how far to move forward and choose which class label to give

State: combination of high-level representation and

assists the agent to make a better decision

Reward: a numerical feedback for each action perfor

Policy learning: a multilayer perceptron to model the

Temporal Convolutional Network TCN)

Input: possible set of states

Layer: policy

Output: possible set of actions

Reinforcement Learning Strategies for Clinical Trials in Nonsmall Cell Lung Cancer. Biometrics

I B S Biometrics JOURNAL OF THE INTERNATIONAL BIOMETRIC SOCIETY
Explore this journal >

Reinforcement Learning Strategies for Clinical Trials in Nonsmall Cell Lung Cancer

Yufan Zhao, Donglin Zeng, Mark A. Socinski, Michael R. Kosorok

First published: 8 March 2011 Full publication history
DOI: 10.1111/j.1541-0420.2011.01572.x View/save citation

Cited by (CrossRef): 30 articles Check for updates | Citation tools ▾

Abstract

Summary Typical regimens for advanced metastatic stage IIIB/IV nonsmall cell lung cancer (NSCLC) consist of multiple lines of treatment. We present an adaptive reinforcement learning approach to discover optimal individualized treatment regimens from a specially designed clinical trial (a "clinical reinforcement trial") of an experimental treatment for patients with advanced NSCLC who have not been treated previously with systemic therapy. In addition to the complexity of the problem of selecting optimal compounds for first- and second-line treatments based on prognostic factors, another primary goal is to determine the optimal time to initiate second-line

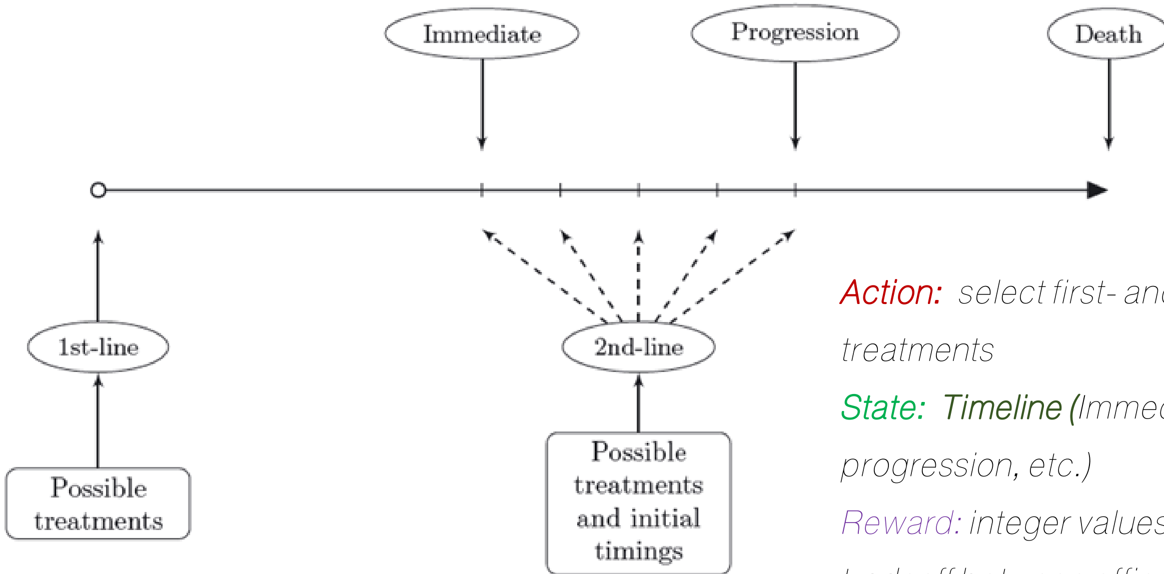
Reinforcement Learning Strategies for Clinical Trials in Nonsmall Cell Lung Cancer. Biometrics

Reinforcement Learning

Target variable not available

Control

Treatment plan and therapy options for an advanced non-small cell lung cancer



Action: select first- and second-line treatments

State: Timeline (Immediate, Disease progression, etc.)

Reward: integer values to assess the tradeoff between efficacy and toxicity

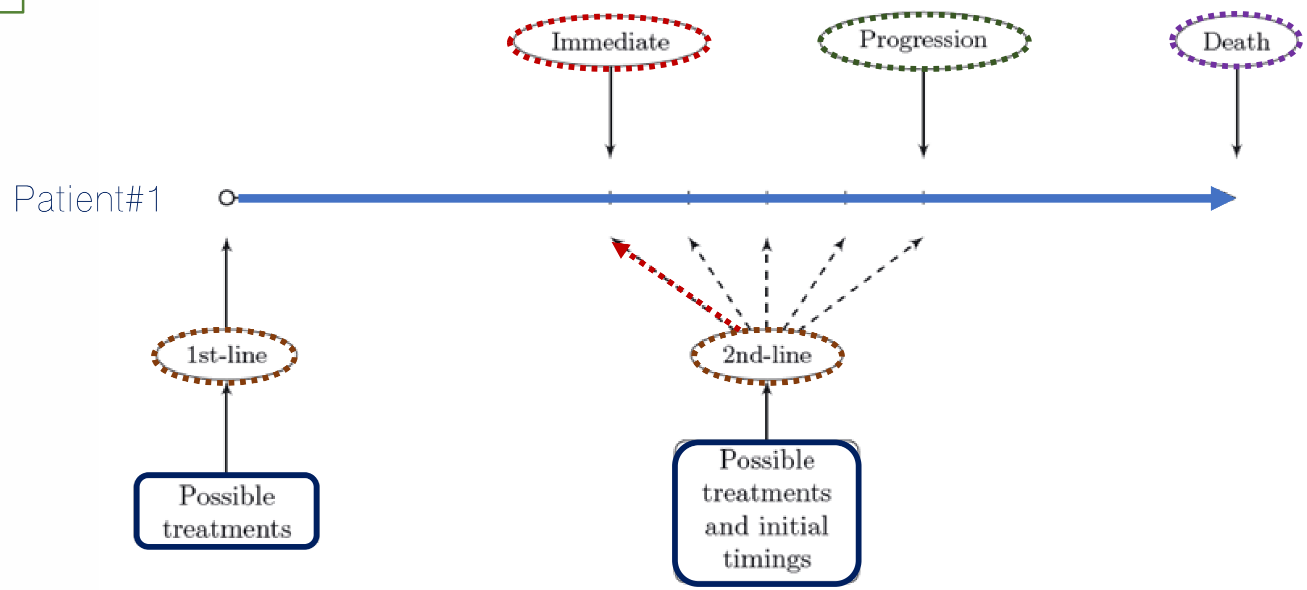
Reinforcement Learning Strategies for Clinical Trials in Nonsmall Cell Lung Cancer. Biometrics

Reinforcement Learning

Target variable not available

Control

Treatment plan and therapy options for an advanced non-small cell lung cancer



Round 1

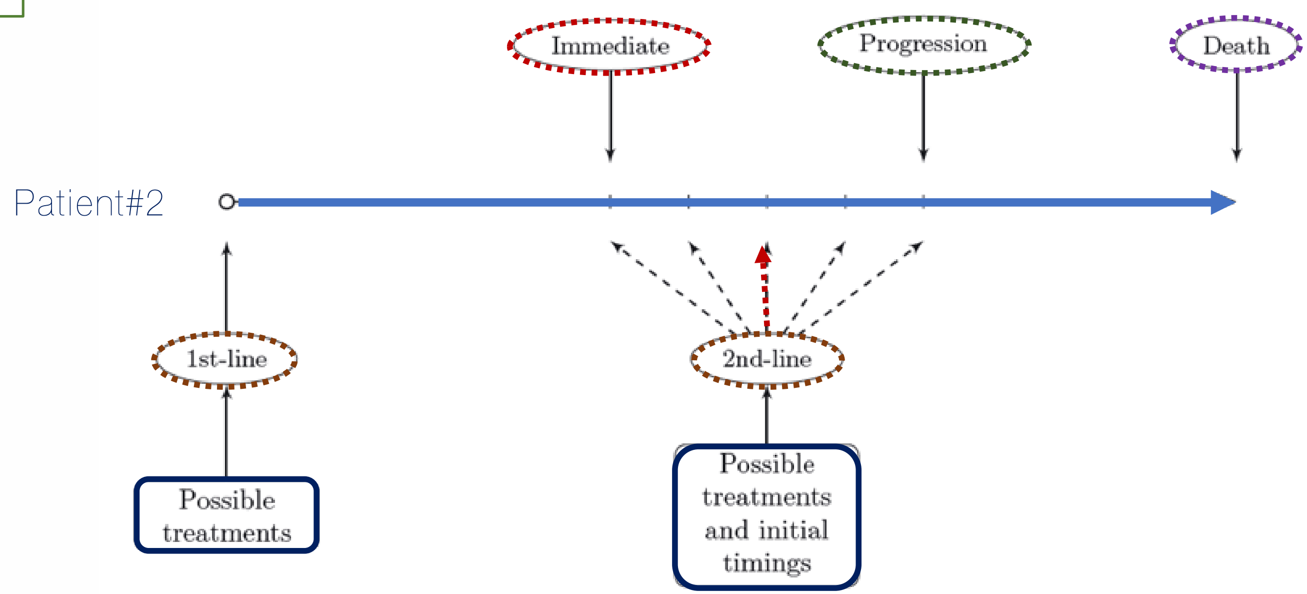
Reinforcement Learning Strategies for Clinical Trials in Nonsmall Cell Lung Cancer. Biometrics

Reinforcement Learning

Target variable not available

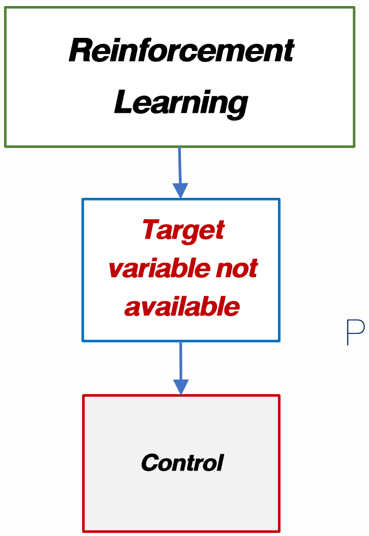
Control

Treatment plan and therapy options for an advanced non-small cell lung cancer

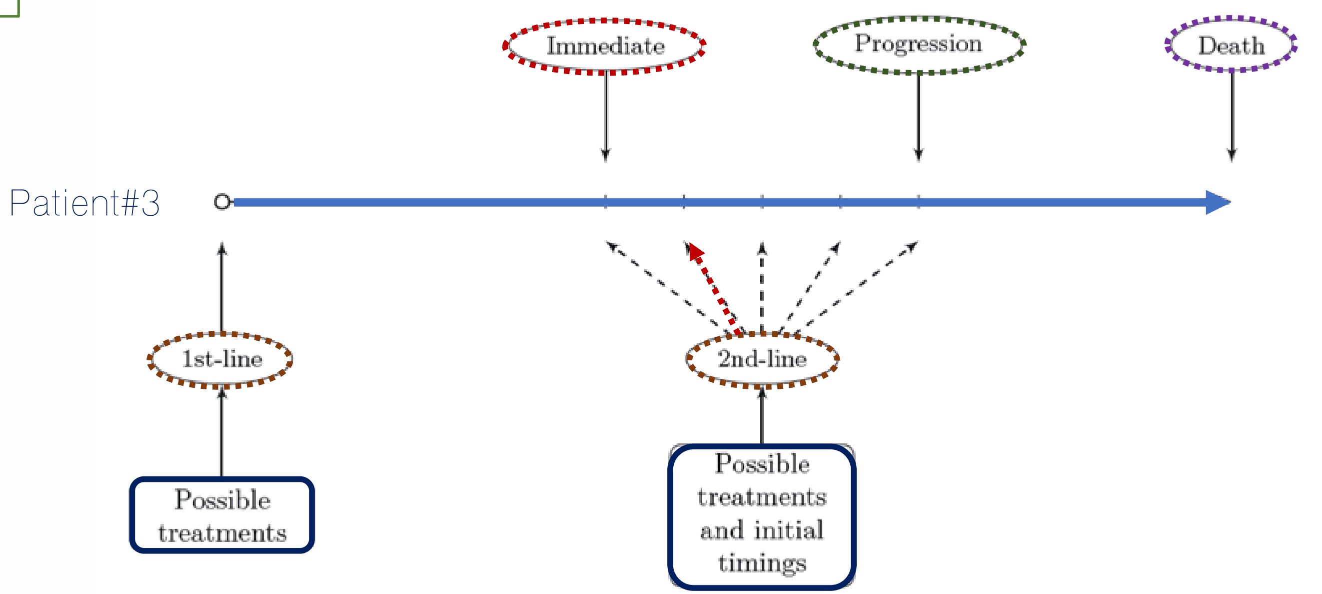


Round 2- Round n

Reinforcement Learning Strategies for Clinical Trials in Nonsmall Cell Lung Cancer. Biometrics

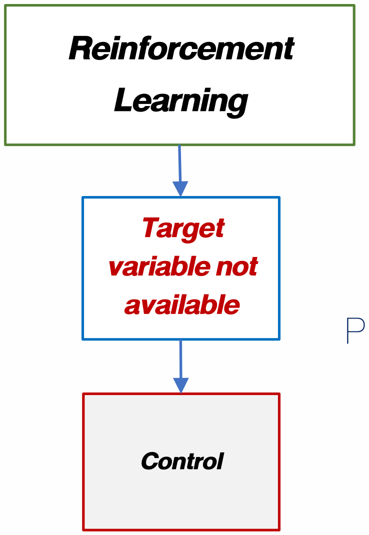


Treatment plan and therapy options for an advanced non-small cell lung cancer

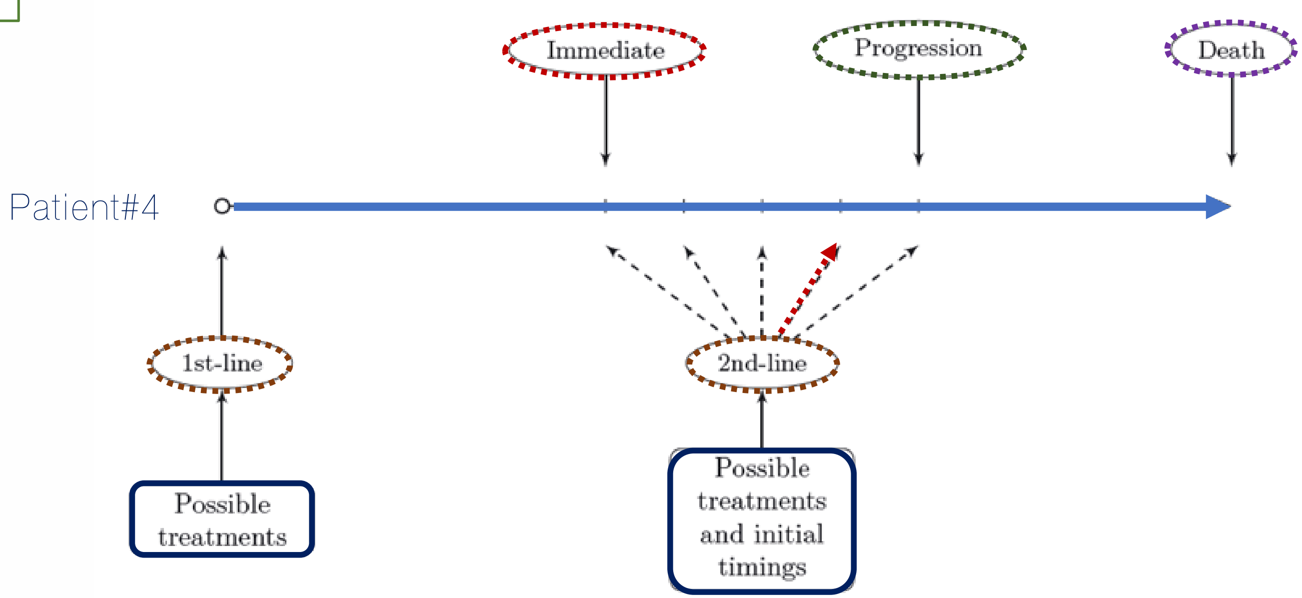


Round 2- Round n

Reinforcement Learning Strategies for Clinical Trials in Nonsmall Cell Lung Cancer. Biometrics



Treatment plan and therapy options for an advanced non-small cell lung cancer



Round 2- Round n

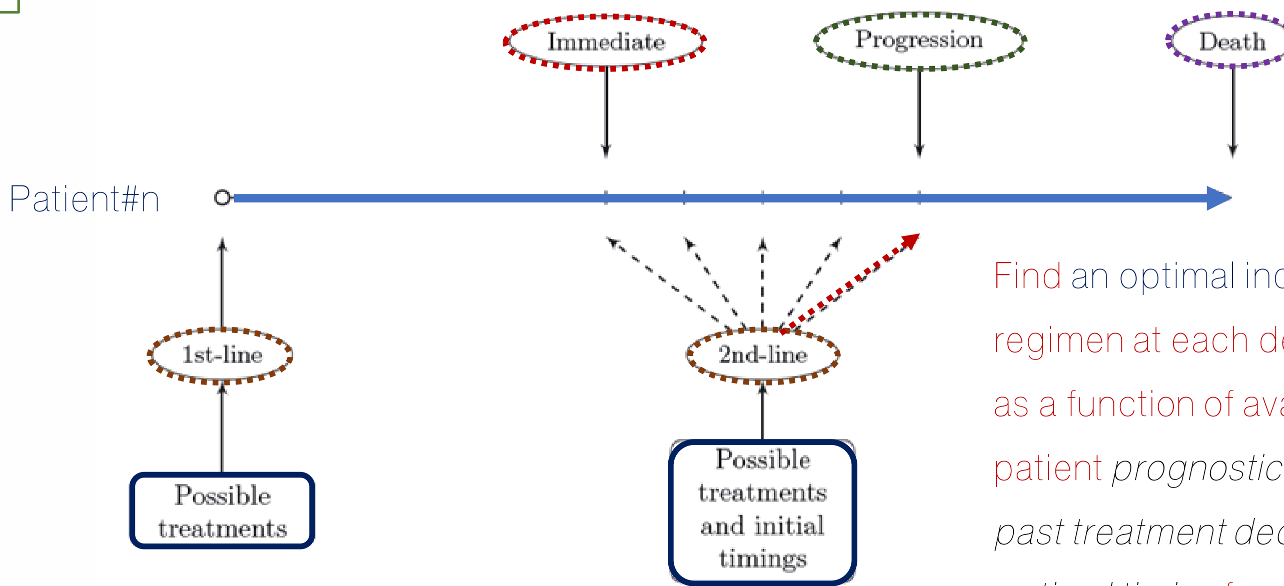
Reinforcement Learning Strategies for Clinical Trials in Nonsmall Cell Lung Cancer. Biometrics

Reinforcement Learning

Target variable not available

Control

Treatment plan and therapy options for an advanced non-small cell lung cancer



Round 2- Round n

Find an optimal individualized regimen at each decision time as a function of available patient *prognostic factors*, *past treatment decisions*, and *optimal timing* for maximizing the average survival time

Pruning an ensemble of classifiers via reinforcement learning

Neurocomputing 72 (2009) 1900–1909

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

ELSEVIER

Pruning an ensemble of classifiers via reinforcement learning

Ioannis Partalas *, Grigorios Tsoumakas, Ioannis Vlahavas

Department of Informatics, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece

ARTICLE INFO

Article history:
Received 13 July 2007
Received in revised form
7 April 2008
Accepted 20 June 2008
Communicated by A. Suarez
Available online 19 July 2008

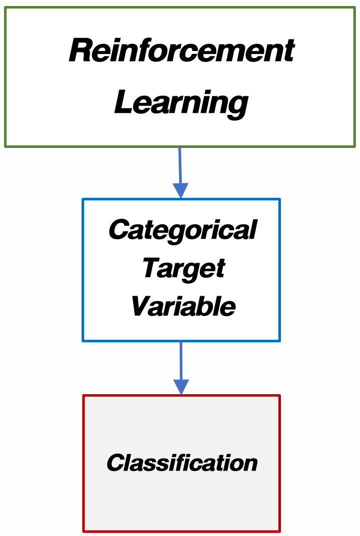
Keywords:
Reinforcement learning
Ensemble selection

ABSTRACT

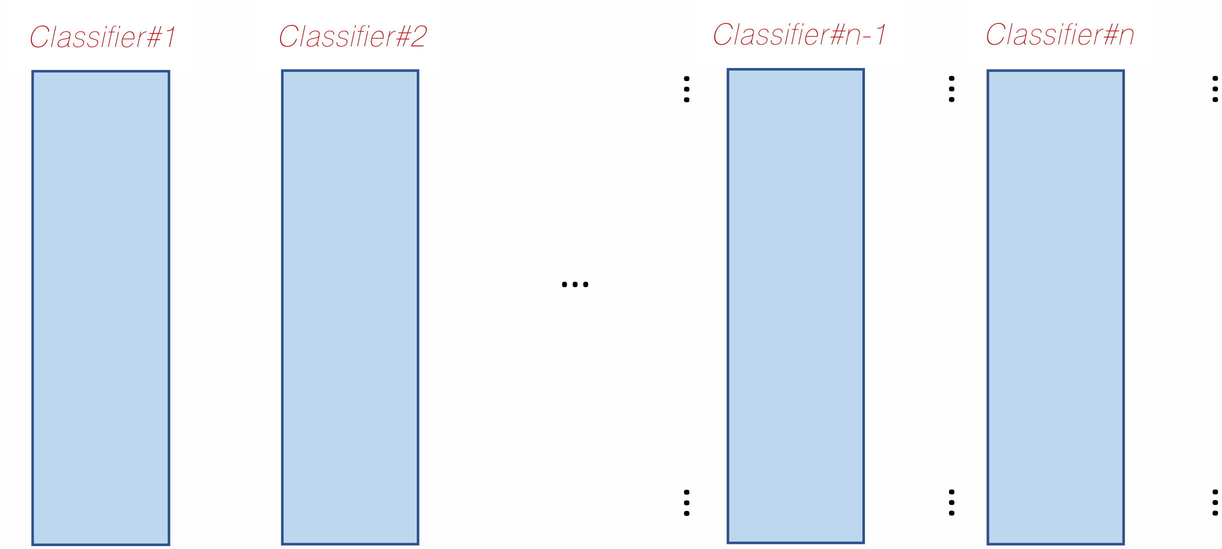
This paper studies the problem of pruning an ensemble of classifiers from a reinforcement learning perspective. It contributes a new pruning approach that uses the Q-learning algorithm in order to approximate an optimal policy of choosing whether to include or exclude each classifier from the ensemble. Extensive experimental comparisons of the proposed approach against state-of-the-art pruning and combination methods show very promising results. Additionally, we present an extension that allows the improvement of the solutions returned by the proposed approach over time, which is very useful in certain performance-critical domains.

© 2008 Elsevier B.V. All rights reserved.

Pruning an ensemble of classifiers via reinforcement learning

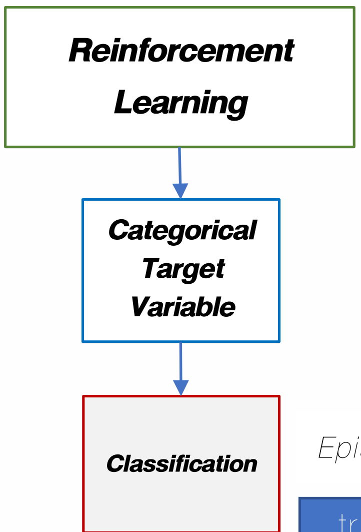


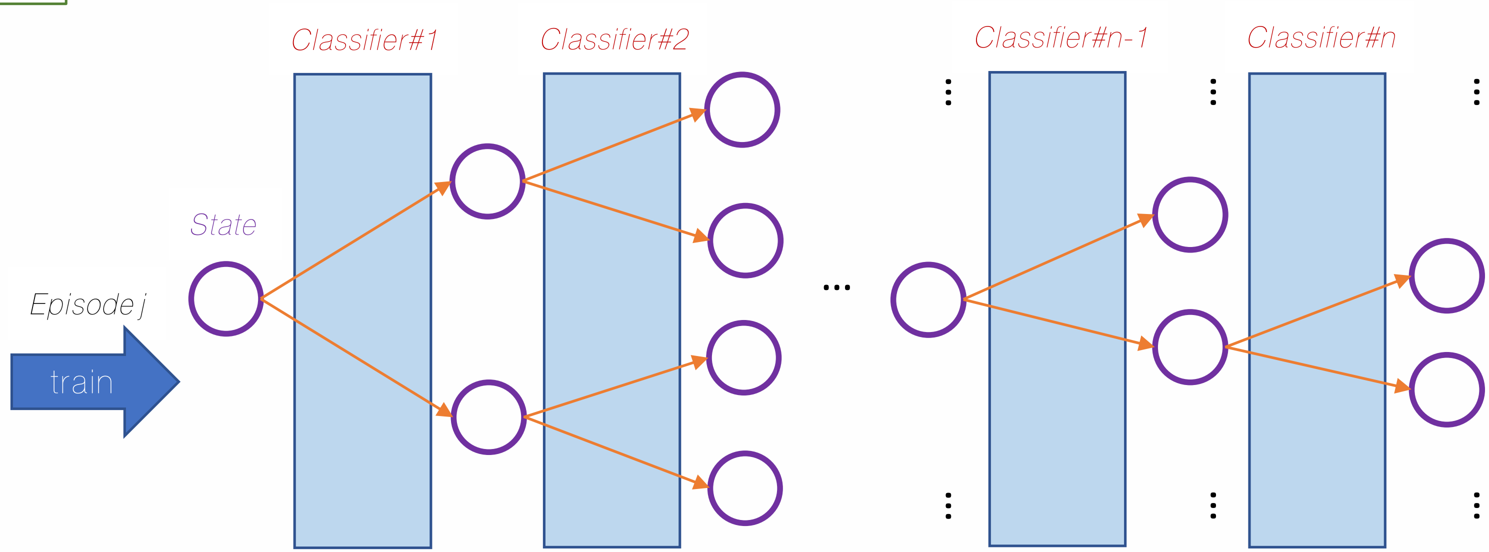
use an agent to prune the ensemble of classifiers and build the sets of good classifiers that based on long-term reward maximization



Pruning an ensemble of classifiers via reinforcement learning

use an agent to prune the ensemble of classifiers and build the sets of good classifiers that based on long-term reward maximization

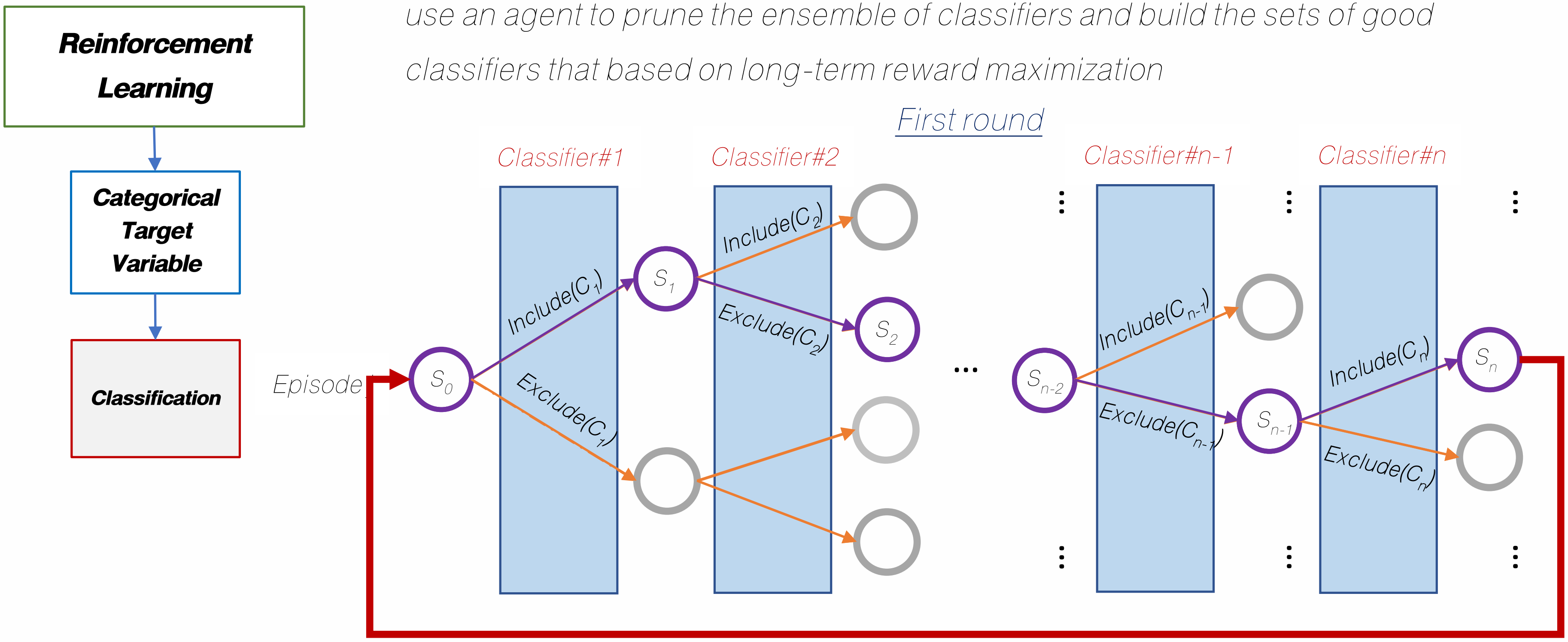




Mahidol University
Faculty of Medicine Ramathibodi Hospital
Department of Clinical Epidemiology and Biostatistics

Pruning an ensemble of classifiers via reinforcement learning

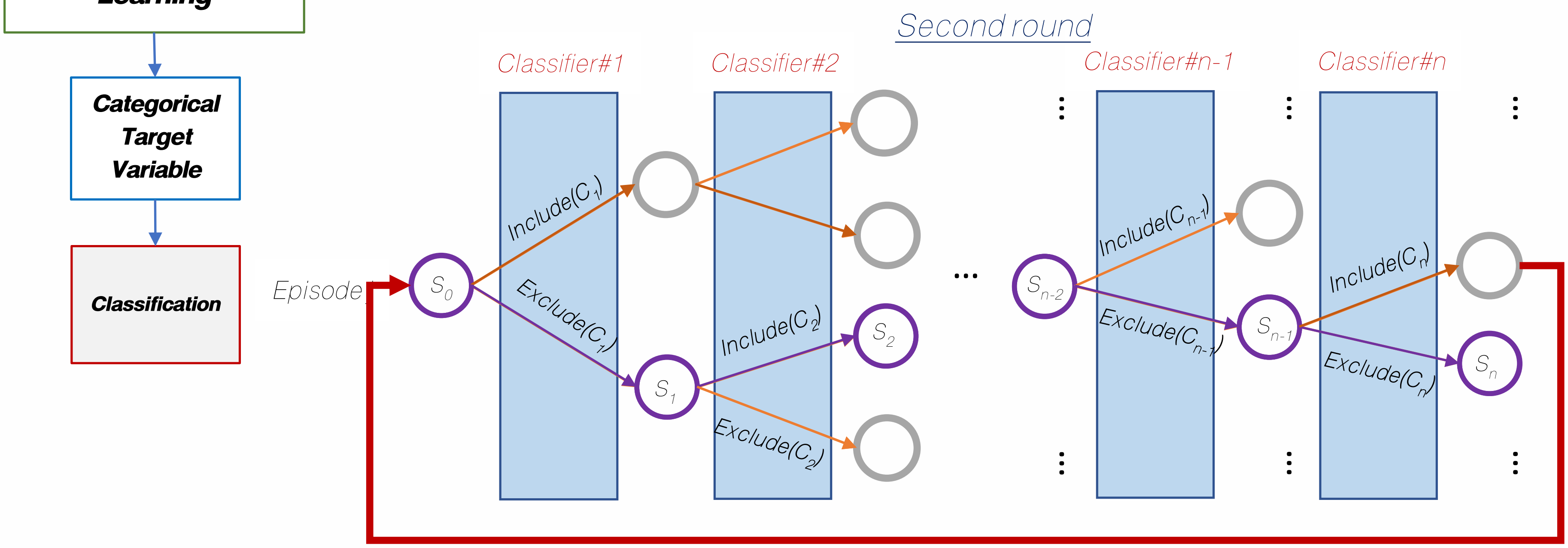
use an agent to prune the ensemble of classifiers and build the sets of good classifiers that based on long-term reward maximization



long-term reward (overall performance) for set of states i^{th} of episode j^{th}

Pruning an ensemble of classifiers via reinforcement learning

use an agent to prune the ensemble of classifiers and build the sets of good classifiers that based on long-term reward maximization



long-term reward for set of states $i+1^{th}$ of episode j^{th}

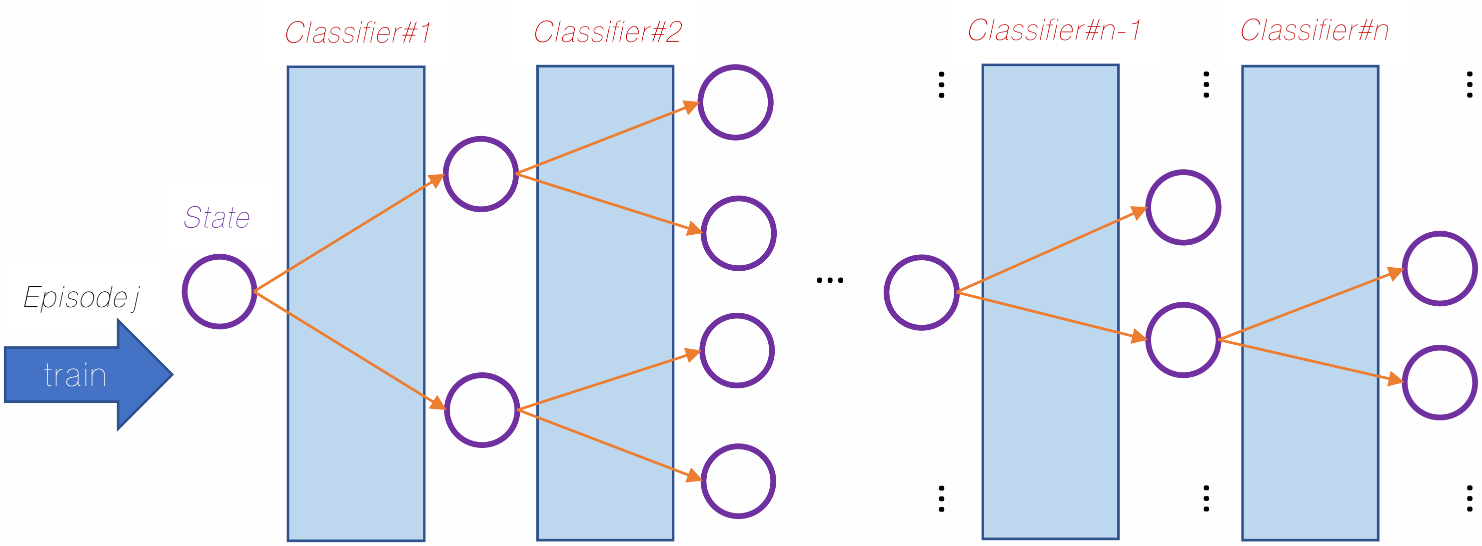
Pruning an ensemble of classifiers via reinforcement learning

Reinforcement Learning

Categorical Target Variable

Classification

use an agent to prune the ensemble of classifiers and build the sets of good classifiers that based on long-term reward maximization



- *Repeat for each episode until reach the iteration condition,*
- *This offers the advantage of requiring less episodes to train the agent*

Mahidol University
Faculty of Medicine Ramathibodi Hospital
Department of Clinical Epidemiology and Biostatistics

References

1. Stember, Joseph & Shalu, Hrithwik. (2020). Title: Deep reinforcement learning to detect brain lesions on MRI: a proof-of-concept application of reinforcement learning to medical images. 10.13140/RG.2.2.20025.80487.
2. Peng, Xuefeng & Ding, Yi & Wihl, David & Gottesman, Omer & Komorowski, Matthieu & Lehman, Li-wei & Ross, Andrew & Faisal, Aldo & Doshi velez, Finale. (2019). Improving Sepsis Treatment Strategies by Combining Deep and Kernel-Based Reinforcement Learning.
3. N Abe, N Verma, C Apte, R Schroko, Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, Pages 767-772 , 2004.
4. Ioannis Partalas, Grigorios Tsoumakas and Ioannis Vlahavas, Pruning an ensemble of classifiers via reinforcement earning, Neurocomputing, vol. 72 pp.1900–1909, 2009.
5. Yufan Zhao, Donglin Zeng, Mark A. Socinski and Michael R. Kosorok , Reinforcement Learning Strategies for Clinical Trials in Nonsmall Cell Lung Cancer. Biometrics. Vol. 67. December 2011.