ADVANCING BEHAVIORAL SCIENCE THROUGH AI + DIGITAL HEALTH

RL-Driven Pain Care Recommendations – Dr. Pratik Gajane

MICHIGAN INSTITUTE FOR CLINICAL & HEALTH RESEARCH
UNIVERSITY OF MICHIGAN

May 21, 2024

BRISP
Behavioral Research Innovation and Support Program

Overview

- Reinforcement learning (RL) to deliver personalized CBT for chronic pain.

Overview

- Reinforcement learning (RL) to deliver personalized CBT for chronic pain.
- In this presentation, we will see
  - how RL-based results display disparities w.r.t. *sensitive* attributes (gender, race), and
  - how to avoid such disparities.

∗ University of Michigan, Ann Arbor, USA.

† Eindhoven University of Technology, Eindhoven, the Netherlands.

Overview

- Reinforcement learning (RL) to deliver personalized CBT for chronic pain.
- In this presentation, we will see
  - how RL-based results display disparities w.r.t. *sensitive* attributes (gender, race), and
  - how to avoid such disparities.
- Latest version of this work : Investigating Gender Fairness in Machine Learning-driven Personalized Care for Chronic Pain (arxiv.org/abs/2402.19226).

∗ University of Michigan, Ann Arbor, USA.

† Eindhoven University of Technology, Eindhoven, the Netherlands.

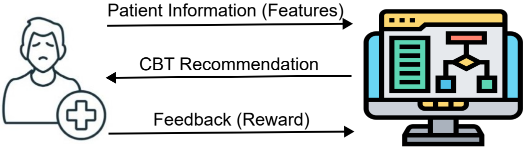
Overview

- Reinforcement learning (RL) to deliver personalized CBT for chronic pain.
- In this presentation, we will see
  - how RL-based results display disparities w.r.t. *sensitive* attributes (gender, race), and
  - how to avoid such disparities.
- Latest version of this work : Investigating Gender Fairness in Machine Learning-driven Personalized Care for Chronic Pain (arxiv.org/abs/2402.19226).
- Collaborators for this work :
  - John D. Piette*,
  - Sean Newman*, and
  - Mykola Pechenizkiy†.

* University of Michigan, Ann Arbor, USA.

† Eindhoven University of Technology, Eindhoven, the Netherlands.

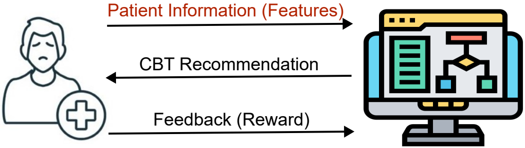
Personalized CBT for Chronic Pain via Reinforcement Learning



CBT Options
Pre-recorded message from a therapist
15-minute in-person session with a therapist
45-minute in-person session with a therapist

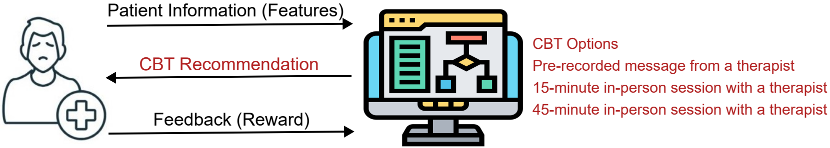
- In each patient interaction, the algorithm

Personalized CBT for Chronic Pain via Reinforcement Learning



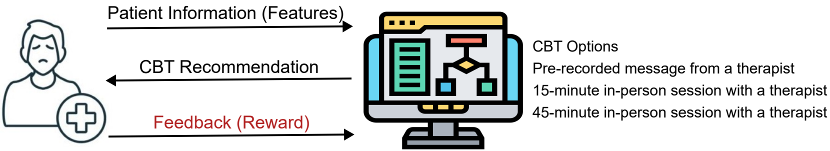
CBT Options
Pre-recorded message from a therapist
15-minute in-person session with a therapist
45-minute in-person session with a therapist

- In each patient interaction, the algorithm
  - Observes relevant patient information (features);



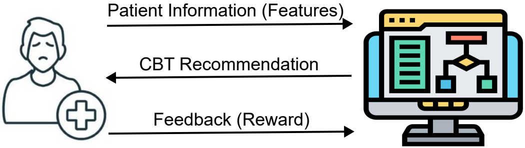
- In each patient interaction, the algorithm
  - Observes relevant patient information (features);
  - Uses features to decide which CBT option to recommend;

Personalized CBT for Chronic Pain via Reinforcement Learning



- In each patient interaction, the algorithm
  - Observes relevant patient information (features);
  - Uses features to decide which CBT option to recommend;
  - Receives a reward (a number expressing the efficacy of recommended CBT option).

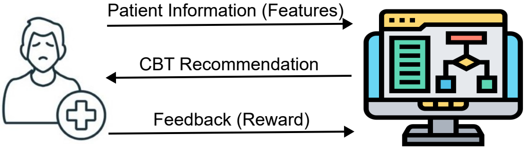
Personalized CBT for Chronic Pain via Reinforcement Learning



CBT Options
Pre-recorded message from a therapist
15-minute in-person session with a therapist
45-minute in-person session with a therapist

- In each patient interaction, the algorithm
  - Observes relevant patient information (features);
  - Uses features to decide which CBT option to recommend;
  - Receives a reward (a number expressing the efficacy of recommended CBT option).
- Earlier patient interactions (features, recommendation, reward) used to select recommendations in the subsequent interactions.

Personalized CBT for Chronic Pain via Reinforcement Learning



CBT Options
Pre-recorded message from a therapist
15-minute in-person session with a therapist
45-minute in-person session with a therapist

- In each patient interaction, the algorithm
  - Observes relevant patient information (features);
  - Uses features to decide which CBT option to recommend;
  - Receives a reward (a number expressing the efficacy of recommended CBT option).
- Earlier patient interactions (features, recommendation, reward) used to select recommendations in the subsequent interactions.
- Performance criterion : Utility = Average received reward.
  High utility reflects effective CBT recommendations.

RL-based CBT Recommendations: Advantages and Concerns

- A recent clinical trial [1] showed that RL-based CBT treatments led to
  - ✓ high utility (leading to improved patient outcomes) and,
  - effective allocation of scarce clinical resources.

[1] Piette et al. Patient-centered pain care using artificial intelligence and mobile health tools: A randomized comparative effectiveness trial, 182(9):975, September 2022a.

RL-based CBT Recommendations: Advantages and Concerns

- A recent clinical trial [1] showed that RL-based CBT treatments led to
  - ✓ high utility (leading to improved patient outcomes) and,
  - effective allocation of scarce clinical resources.

✗ (RL) Algorithms can introduce/amplify disparities w.r.t. sensitive attributes (gender, race, etc)!

[1] Piette et al. Patient-centered pain care using artificial intelligence and mobile health tools: A randomized comparative effectiveness trial, 182(9):975, September 2022a.

RL-based CBT Recommendations: Advantages and Concerns

- A recent clinical trial [1] showed that RL-based CBT treatments led to
  - ✓ high utility (leading to improved patient outcomes) and,
  - effective allocation of scarce clinical resources.

✗ (RL) Algorithms can introduce/amplify disparities w.r.t. sensitive attributes (gender, race, etc)!

- We investigated gender disparities in the utility of the RL algorithm for 50000 patient interactions based on [2].

[1] Piette et al. Patient-centered pain care using artificial intelligence and mobile health tools: A randomized comparative effectiveness trial, 182(9):975, September 2022a.

[2] Dataset #2 for Piette et al.: Data for a Reinforcement Learning Intervention to Treat Chronic Pain.

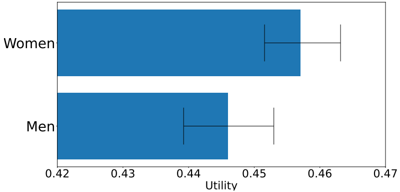
RL-based CBT Recommendations: Advantages and Concerns

- A recent clinical trial [1] showed that RL-based CBT treatments led to
  - ✓ high utility (leading to improved patient outcomes) and,
  - effective allocation of scarce clinical resources.

✗ (RL) Algorithms can introduce/amplify disparities w.r.t. sensitive attributes (gender, race, etc)!

- We investigated gender disparities in the utility of the RL algorithm for 50000 patient interactions based on [2].

✗ Utility for women $>$ Utility for men, p=0.01, Cohen's d=0.32.



[1] Piette et al. Patient-centered pain care using artificial intelligence and mobile health tools: A randomized comparative effectiveness trial, 182(9):975, September 2022a.

[2] Dataset #2 for Piette et al.: Data for a Reinforcement Learning Intervention to Treat Chronic Pain.

RL-based CBT Recommendations: Advantages and Concerns

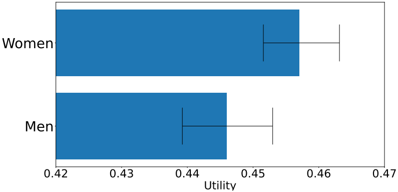
- A recent clinical trial [1] showed that RL-based CBT treatments led to
  - ✓ high utility (leading to improved patient outcomes) and,
  - effective allocation of scarce clinical resources.

✗ (RL) Algorithms can introduce/amplify disparities w.r.t. sensitive attributes (gender, race, etc)!

- We investigated gender disparities in the utility of the RL algorithm for 50000 patient interactions based on [2].

✗ Utility for women $>$ Utility for men, p=0.01, Cohen's d=0.32.

- High utility Yes
  Equity w.r.t. gender No



[1] Piette et al. Patient-centered pain care using artificial intelligence and mobile health tools: A randomized comparative effectiveness trial, 182(9):975, September 2022a.

[2] Dataset #2 for Piette et al.: Data for a Reinforcement Learning Intervention to Treat Chronic Pain.

High Utility and Equity via Feature Selection

- Patient Features

Pain interference 1	Pain interference 2
Session Number	% days this week with steps goal met
CBT skill practice this week	Pain intensity change
Sleep quality this week	Sleep duration this week

High Utility and Equity via Feature Selection

- Patient Features

Pain interference 1	Pain interference 2
Session Number	% days this week with steps goal met
CBT skill practice this week	Pain intensity change
Sleep quality this week	Sleep duration this week

Feature Selection : Use a subset of features.

High Utility and Equity via Feature Selection

- Patient Features

Pain interference 1	Pain interference 2
Session Number	% days this week with steps goal met
CBT skill practice this week	Pain intensity change
Sleep quality this week	Sleep duration this week

Feature Selection : Use a subset of features.

- We evaluated the RL algorithm with 22 feature combinations, and almost all resulted in gender disparities ($p < 0.05$ and Cohen's d around 0.3).

High Utility and Equity via Feature Selection

- Patient Features

~~Pain interference 1~~	~~Pain interference 2~~
Session Number	% days this week with steps goal met
CBT skill practice this week	Pain intensity change
Sleep quality this week	Sleep duration this week

Optimal Features

Feature Selection : Use a subset of features.

- We evaluated the RL algorithm with 22 feature combinations, and almost all resulted in gender disparities ($p < 0.05$ and Cohen's d around 0.3).
- Results using Optimal Features
  - Utility for men : Highest among all combinations
    Utility for women : Highest among all combinations

High Utility and Equity via Feature Selection

- Patient Features

~~Pain interference 1~~	~~Pain interference 2~~
Session Number	% days this week with steps goal met
CBT skill practice this week	Pain intensity change
Sleep quality this week	Sleep duration this week

Optimal Features

Feature Selection : Use a subset of features.

- We evaluated the RL algorithm with 22 feature combinations, and almost all resulted in gender disparities ($p < 0.05$ and Cohen's d around 0.3).
- Results using Optimal Features
  - ✓ Utility for men : Highest among all combinations
    Utility for women : Highest among all combinations
  - ✓ Utility for women $\overset{\text{not statistically different}}{\approx}$ utility for men.

High Utility and Equity via Feature Selection

- Patient Features

~~Pain interference 1~~	~~Pain interference 2~~
Session Number	% days this week with steps goal met
CBT skill practice this week	Pain intensity change
Sleep quality this week	Sleep duration this week

Optimal Features

Feature Selection : Use a subset of features.

- We evaluated the RL algorithm with 22 feature combinations, and almost all resulted in gender disparities ($p < 0.05$ and Cohen's d around 0.3).
- Results using Optimal Features
  - ✓ Utility for men : Highest among all combinations
    Utility for women : Highest among all combinations
  - ✓ Utility for women $\overset{\text{not statistically different}}{\approx}$ utility for men.
- If the RL algorithm makes recommendations based on Optimal Features
  High utility Yes
  Equity w.r.t. gender Yes

Objective: Recommend CBT treatments using Optimal Features.

Our RL Algorithm with Feature Selection for Utility and Equity

Objective: Recommend CBT treatments using Optimal Features.

❷ The identity of Optimal Features may be unknown.

Our RL Algorithm with Feature Selection for Utility and Equity

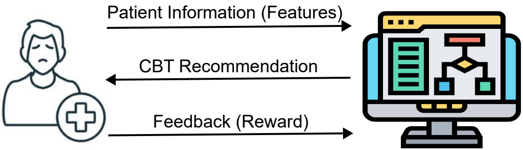
Objective: Recommend CBT treatments using Optimal Features.

- The identity of Optimal Features may be unknown.
- Optimal Features may differ across patient populations.

Our RL Algorithm with Feature Selection for Utility and Equity

Objective: Recommend CBT treatments using Optimal Features.

- ❓ The identity of Optimal Features may be unknown.
- ❓ Optimal Features may differ across patient populations.

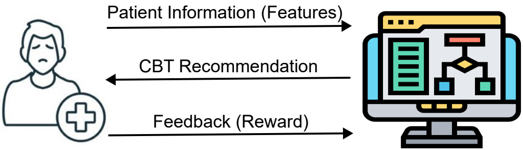


- In each patient interaction, the algorithm
  - Observes features;
  - **Feature selection : Selects a subset of features**;

Our RL Algorithm with Feature Selection for Utility and Equity

Objective: Recommend CBT treatments using Optimal Features.

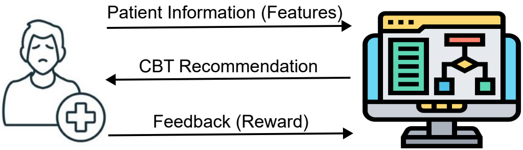
- ❓ The identity of Optimal Features may be unknown.
- ❓ Optimal Features may differ across patient populations.



- In each patient interaction, the algorithm
  - Observes features;
  - **Feature selection : Selects a subset of features**;
  - Uses **selected features** to decide which treatment to recommend;

Objective: Recommend CBT treatments using Optimal Features.

- ❓ The identity of Optimal Features may be unknown.
- ❓ Optimal Features may differ across patient populations.



- In each patient interaction, the algorithm
  - Observes features;
  - **Feature selection : Selects a subset of features**;
  - Uses **selected features** to decide which treatment to recommend;
  - Receives a reward.
- Earlier patient interactions (features, recommendation, reward) used to select subsequent recommendations and for **feature selection**.

Results for Our RL algorithm with Feature Selection

- Evaluated our RL algorithm on 50000 patient interactions based on [2].
- Our algorithm used Optimal Features to make CBT recommendations for 86% of patient interactions.

[2] Dataset #2 for Piette et al.: Data for a Reinforcement Learning Intervention to Treat Chronic Pain.

Results for Our RL algorithm with Feature Selection

- Evaluated our RL algorithm on 50000 patient interactions based on [2].
- Our algorithm used Optimal Features to make CBT recommendations for 86% of patient interactions.
- Relatively infrequent use of Optimal Features initially (*Cold start*).
  $\Rightarrow$ Initial patient interactions possibly receiving poor recommendations.

[2] Dataset #2 for Piette et al.: Data for a Reinforcement Learning Intervention to Treat Chronic Pain.

Results for Our RL algorithm with Feature Selection

- Evaluated our RL algorithm on 50000 patient interactions based on [2].
- Our algorithm used Optimal Features to make CBT recommendations for 86% of patient interactions.
- Relatively infrequent use of Optimal Features initially (*Cold start*).
  $\Rightarrow$ Initial patient interactions possibly receiving poor recommendations.

Input clinicians' domain knowledge about which features are likely to be optimal.

[2] Dataset #2 for Piette et al.: Data for a Reinforcement Learning Intervention to Treat Chronic Pain.

Results for Our RL algorithm with Feature Selection

- Evaluated our RL algorithm on 50000 patient interactions based on [2].
- Our algorithm used Optimal Features to make CBT recommendations for 86% of patient interactions.
- Relatively infrequent use of Optimal Features initially (*Cold start*).
  $\Rightarrow$ Initial patient interactions possibly receiving poor recommendations.

Input clinicians' domain knowledge about which features are likely to be optimal.

- This led to 17% improvement in our algorithm's frequency of using Optimal Features initially (first decile of interactions).

[2] Dataset #2 for Piette et al.: Data for a Reinforcement Learning Intervention to Treat Chronic Pain.

Advantages of our RL algorithm with Feature Selection

- Can provide CBT recommendations using Optimal Features in real-time.

Advantages of our RL algorithm with Feature Selection

- Can provide CBT recommendations using Optimal Features in real-time.
- Data-driven and automated

Advantages of our RL algorithm with Feature Selection

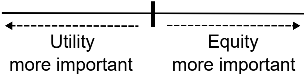
- Can provide CBT recommendations using Optimal Features in real-time.
- Data-driven and automated (with some clinician control).

Advantages of our RL algorithm with Feature Selection

- Can provide CBT recommendations using Optimal Features in real-time.
- Data-driven and automated (with some clinician control).
- Adaptive to varying Optimal Features across patient populations.

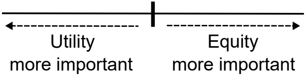
Advantages of our RL algorithm with Feature Selection

- Can provide CBT recommendations using Optimal Features in real-time.
- Data-driven and automated (with some clinician control).
- Adaptive to varying Optimal Features across patient populations.
- Adjustable goal : Can make CBT recommendations targeting —
  - Utility and Equity



Advantages of our RL algorithm with Feature Selection

- Can provide CBT recommendations using Optimal Features in real-time.
- Data-driven and automated (with some clinician control).
- Adaptive to varying Optimal Features across patient populations.
- Adjustable goal: Can make CBT recommendations targeting —
  - Utility and Equity



  - Equity w.r.t. other attributes.

Takeaways

- Uncritical use of patient data with RL algorithms
  - ✓ high utility, but
  - ✗ disparities w.r.t gender (and other sensitive attributes).

Takeaways

- Uncritical use of patient data with RL algorithms
  - ✓ high utility, but
  - ✗ disparities w.r.t gender (and other sensitive attributes).
- Critical use of patient data with RL algorithms
  - ✓ high utility, and
  - ✓ equity w.r.t. gender (and other sensitive attributes).

Takeaways

- Uncritical use of patient data with RL algorithms
  - ✓ high utility, but
  - ✗ disparities w.r.t gender (and other sensitive attributes).
- Critical use of patient data with RL algorithms
  - ✓ high utility, and
  - ✓ equity w.r.t. gender (and other sensitive attributes).

Future Directions

- Patient responses to RL decisions change over time.

Takeaways and Future Directions

Takeaways

- Uncritical use of patient data with RL algorithms
  - ✓ high utility, but
  - ✗ disparities w.r.t gender (and other sensitive attributes).
- Critical use of patient data with RL algorithms
  - ✓ high utility, and
  - ✓ equity w.r.t. gender (and other sensitive attributes).

Future Directions

- Patient responses to RL decisions change over time.
- Learning to defer : Suppose an RL policy, learned on US patient data, is to be deployed in Honduras.

Takeaways

- Uncritical use of patient data with RL algorithms
  - ✓ high utility, but
  - ✗ disparities w.r.t gender (and other sensitive attributes).
- Critical use of patient data with RL algorithms
  - ✓ high utility, and
  - ✓ equity w.r.t. gender (and other sensitive attributes).

Future Directions

- Patient responses to RL decisions change over time.
- Learning to defer : Suppose an RL policy, learned on US patient data, is to be deployed in Honduras.

  When algorithmic recommendations are deemed unreliable/unsafe,
  - the algorithm defers to a human expert;

Takeaways

- Uncritical use of patient data with RL algorithms
  - ✓ high utility, but
  - ✗ disparities w.r.t gender (and other sensitive attributes).
- Critical use of patient data with RL algorithms
  - ✓ high utility, and
  - ✓ equity w.r.t. gender (and other sensitive attributes).

Future Directions

- Patient responses to RL decisions change over time.
- Learning to defer: Suppose an RL policy, learned on US patient data, is to be deployed in Honduras.

  When algorithmic recommendations are deemed unreliable/unsafe,
  - the algorithm defers to a human expert;
  - the human expert makes a reliable recommendation;

Takeaways and Future Directions

Takeaways

- Uncritical use of patient data with RL algorithms
  - ✓ high utility, but
  - ✗ disparities w.r.t gender (and other sensitive attributes).
- Critical use of patient data with RL algorithms
  - ✓ high utility, and
  - ✓ equity w.r.t. gender (and other sensitive attributes).

Future Directions

- Patient responses to RL decisions change over time.
- Learning to defer: Suppose an RL policy, learned on US patient data, is to be deployed in Honduras.

  When algorithmic recommendations are deemed unreliable/unsafe,
  - the algorithm defers to a human expert;
  - the human expert makes a reliable recommendation;
  - the algorithm observes this interaction and learns to make a reliable recommendation in similar future interactions.

Thank you

For more details about this work, see arxiv.org/abs/2402.19226 or my webpage pratikgajane.github.io.

If you'd like to chat, email me at `pratik.gajane@gmail.com`.

MICHR

MICHIGAN INSTITUTE FOR DATA SCIENCE
UNIVERSITY OF MICHIGAN

E-HEALTH AND ARTIFICIAL INTELLIGENCE

BRISP
Behavioral Research Innovation and Support Program