

Federated Reinforcement Learning for Therapeutic Interventions over ICUs with Noisy Labels

Linxiao Cao¹, Yifei Zhu², Haoquan Zhou³, Shilei Tan¹ and Wei Gong¹

1 University of Science and Technology of China

2 Shanghai Jiao Tong University

3 The First Affiliated Hospital of USTC

➤ **Management of routine ICU interventions constitute a major part of intensive care, e.g.:**



➤ **Invasive mechanical ventilation:** the use of mechanical means to assist or replace spontaneous breathing. ^[1]

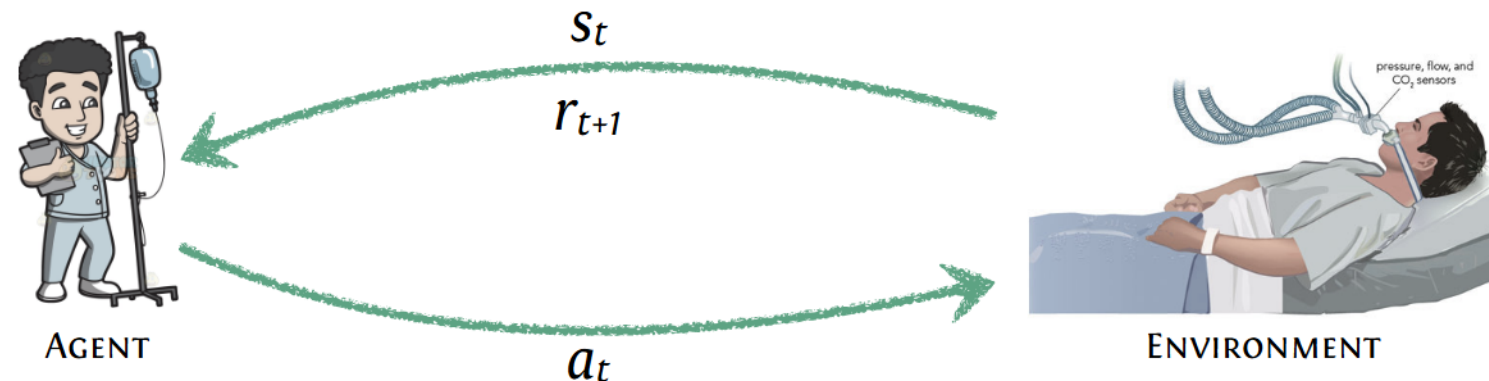
- **40%** of ICU patients are ventilated at any given hour, accounting for **12%** of US hospital costs.
- Typically coupled with sedation to maintain comfort and stability.
- Timely intervention can improve outcomes and reduce costs.
- However, their effects are often poorly understood, especially in heterogeneous patient populations, leading to varying clinical opinions.

[1] A Reinforcement Learning Approach to Weaning of Mechanical Ventilation in Intensive Care Units. UAI 2018.

➤ **Weaning: process of liberation from mechanical ventilation.** [1]

- *Premature, delayed weaning both associated with worse outcomes.*
- We aim to develop **a clinician-in-loop decision support tool** to
 - *alert caregivers when a patient is ready for weaning*
 - *recommend sedation and ventilation settings*

...by modeling this as a Markov Decision Process (MDP).

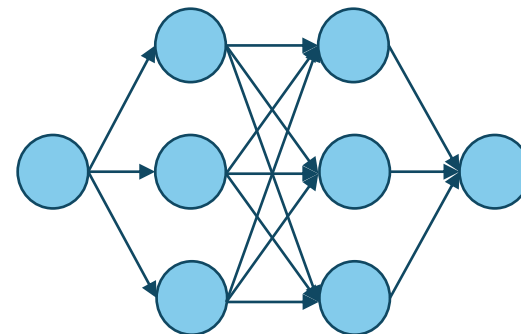


➤ Challenge 1: ICU Data Utilization

- **Limited Data Volume:** Individual hospitals often have limited patient data, which restricts the development of robust predictive models.
- **Data Incompleteness:** The data collected may not be comprehensive, affecting the accuracy and reliability of clinical decisions.



Collaborative Learning



The **sensitivity** of clinical private data
leads to the formation of data silos

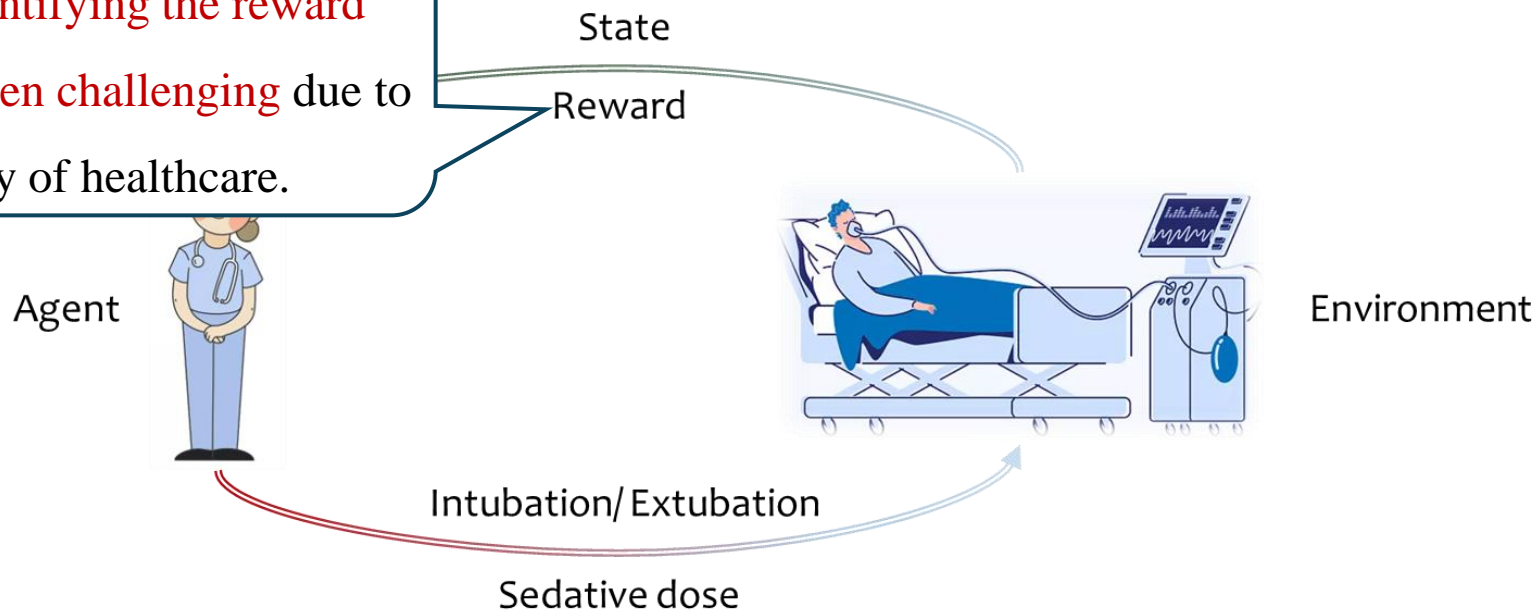
A **data-driven** global clinical
decision-making support tool

How to construct the global model without sharing scattered private clinical raw data?

➤ Challenge 2: Addressing the difficulty of defining a reward function

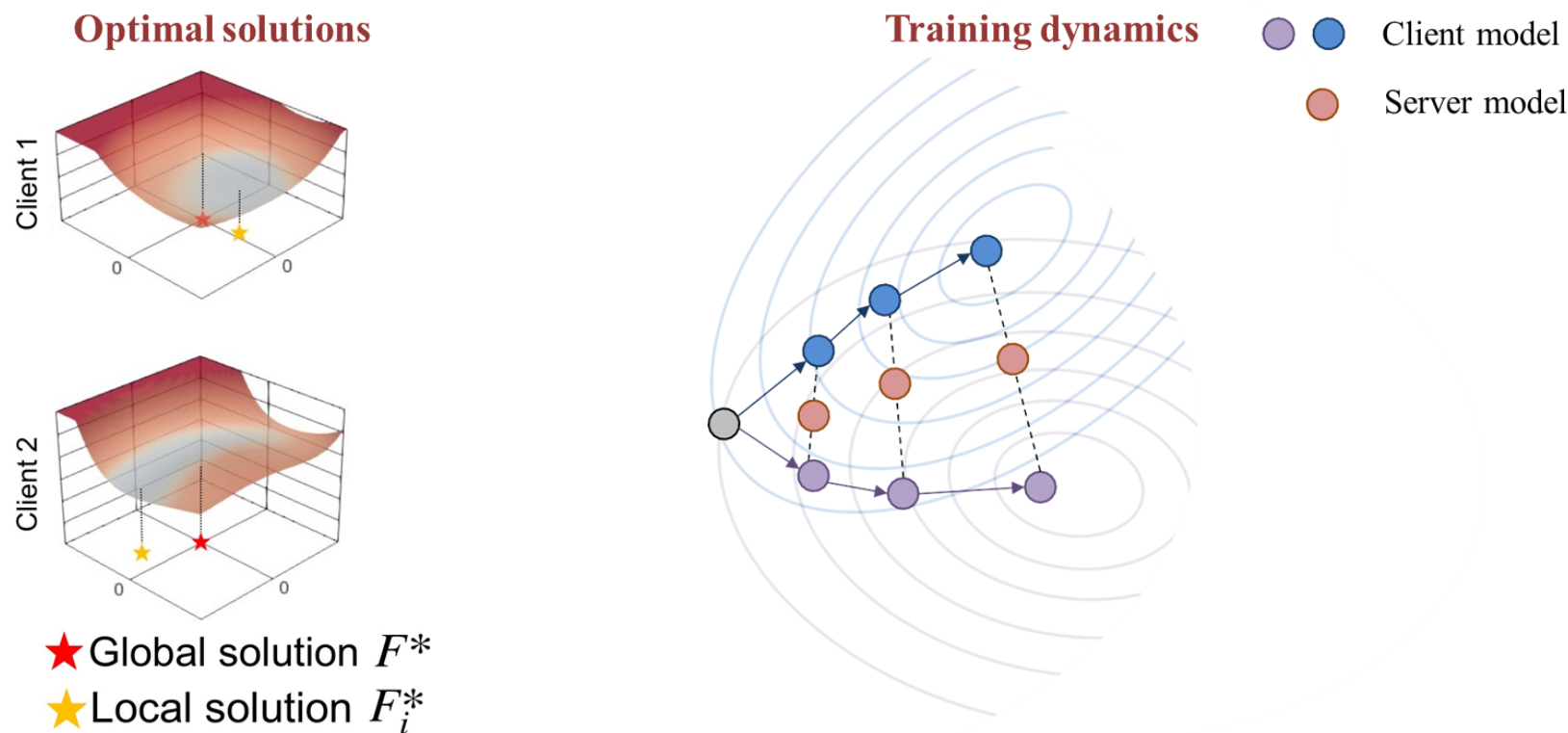
- *In the medical field, many tasks involve sequential decision-making.*
- *These tasks are typically modeled as Markov Decision Processes (MDPs) and tackled with reinforcement learning.*

However, **quantifying the reward function is often challenging** due to the complexity of healthcare.



Inaccurate or arbitrary reward functions may have a negative impact on the policy model.

➤ Challenge 3: Overcoming data heterogeneity among different hospitals

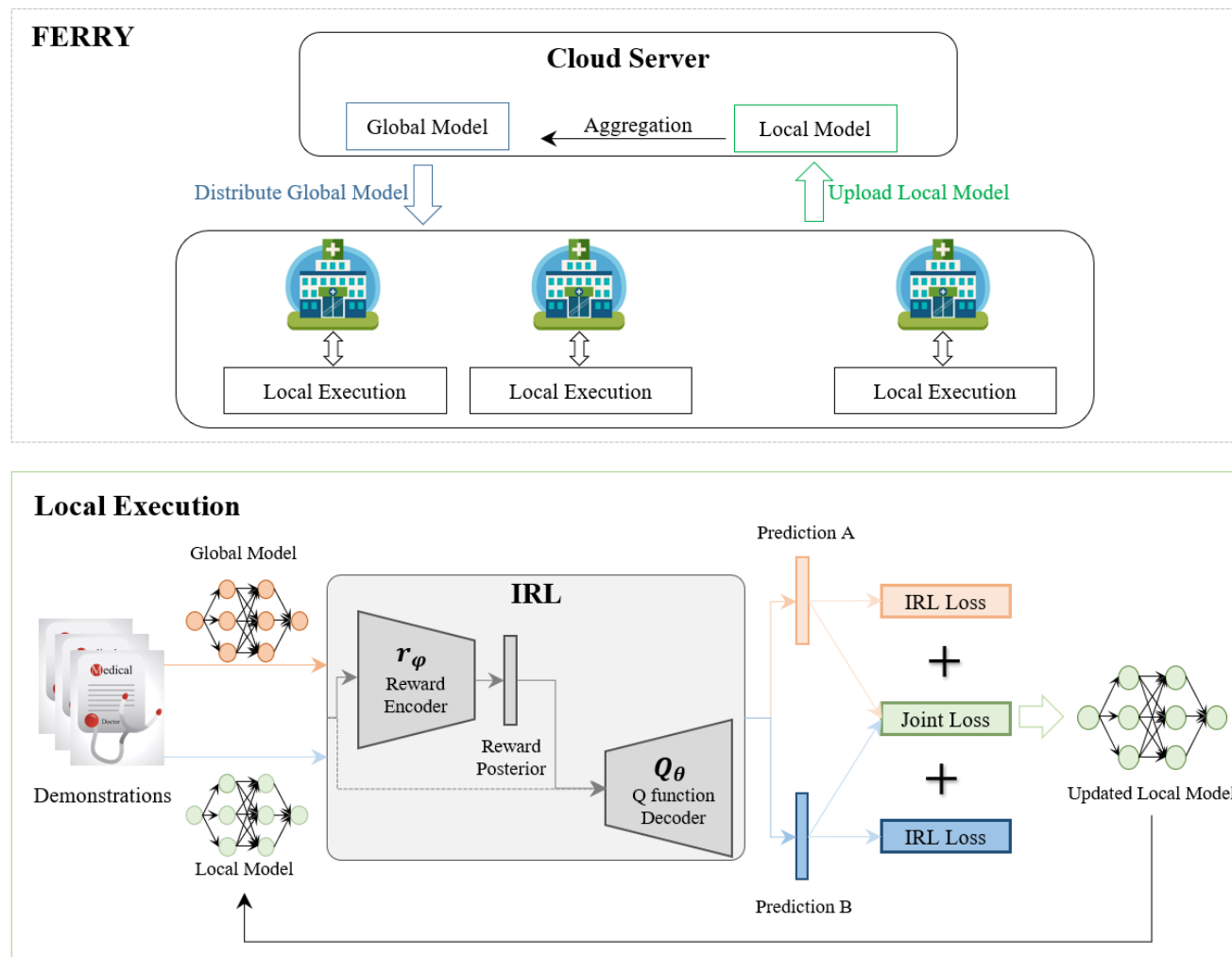


Data heterogeneity can lead to suboptimal performance of global models.^[1]

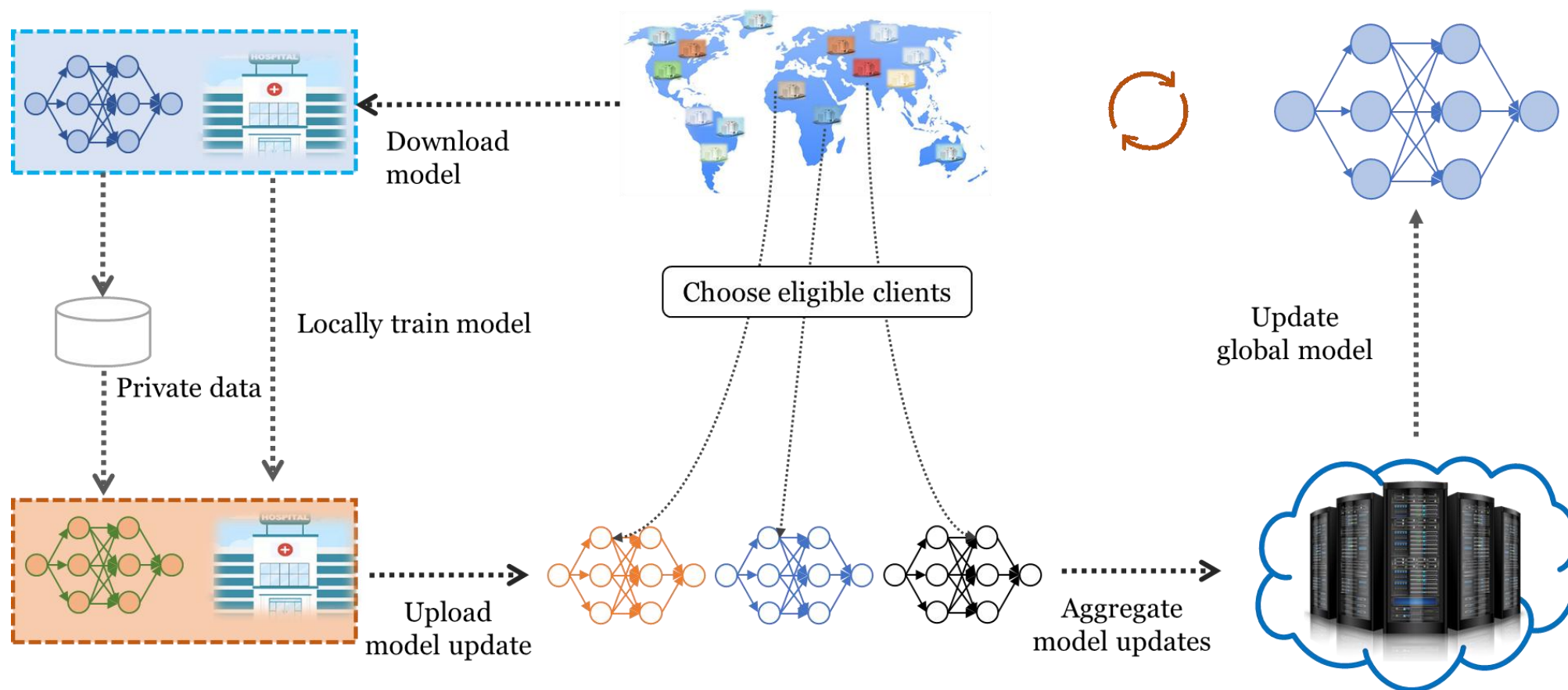
[1] Jiang M, Wang Z, Dou Q. Harmofl: Harmonizing local and global drifts in federated learning on heterogeneous medical images. AAAI 2022.

➤ FERRY: Federated Inverse Reinforcement Learning for Smart ICUs

□ The overview of FERRY

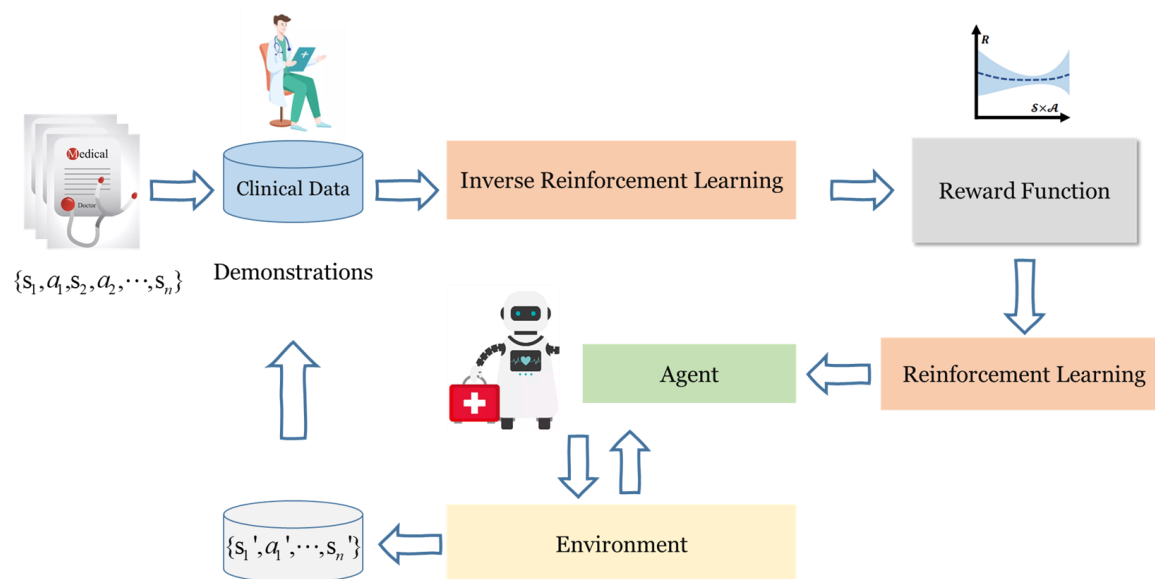


- Introducing federated learning (FL) to break data silos.

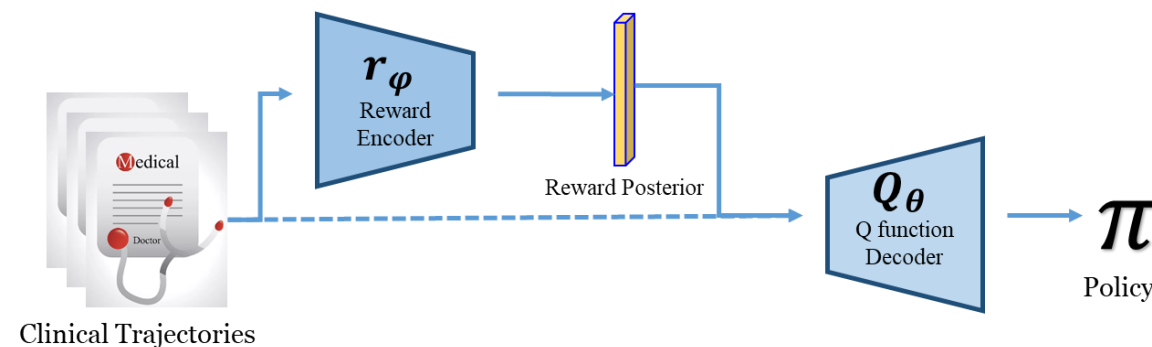


FL ensures data privacy and allows collaborative training using data from other hospitals.

- Using inverse reinforcement learning (IRL) to solve the difficulty of defining reward functions



IRL



AVRIL [1]

IRL can derive suitable reward functions from historical data.

➤ Utilize distributionally robust optimization (DRO) to alleviate the impact of data heterogeneity.

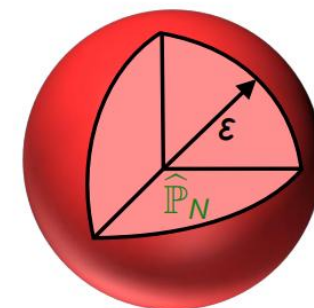
- DRO treats the data distribution of each hospital as an uncertainty factor and employs the approach of minimizing the worst-case under specified uncertainty to train the global model.

$$\min_{\mathbf{w} \in \mathcal{W}} \max_{\boldsymbol{\lambda} \in \Lambda} F(\mathbf{w}, \boldsymbol{\lambda}) := \sum_{i=1}^N \lambda_i f_i(\mathbf{w})$$

- Use the Wasserstein distance to measure the distance between uncertainty sets and probability distributions.

$$\text{Wasserstein metric: } d_W(\mathbb{Q}_1, \mathbb{Q}_2) = \sup_{\text{lip}(f) \leq 1} \mathbb{E}^{\mathbb{Q}_1}[f(\xi)] - \mathbb{E}^{\mathbb{Q}_2}[f(\xi)]$$

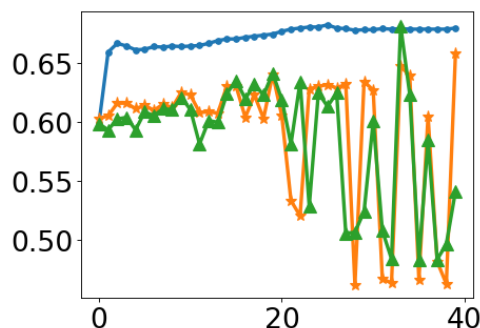
- During each global round, the client with the worst performance is selected for further optimization.



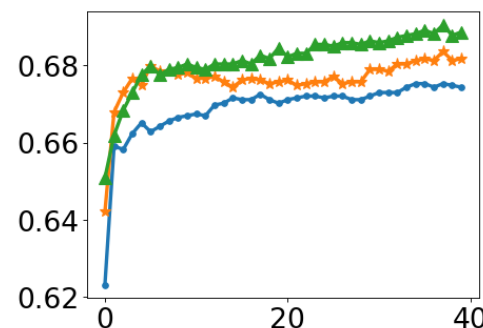
DRO enhances the robustness and generalization capacity of FERRY!

➤ Utilizing joint loss learning to mitigate the impact of noise

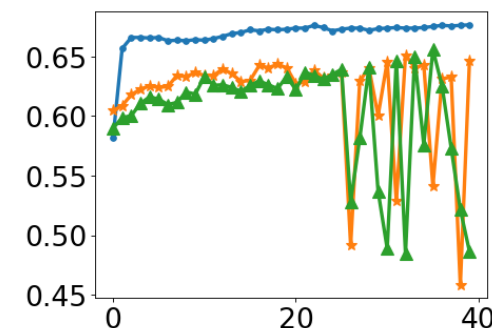
- Previous research demonstrates that DRO can be sensitive to noisy data ^{[1][2]}.
 - DRO uses measurement methods such as norm or divergence in modeling uncertainty distribution, which takes the noise in the uncertainty set into consideration, making DRO sensitive to noise.



Original



No Outliers



Noise Added

Medical data has incorrect labels, and DRO's sensitivity to noise affects model performance!

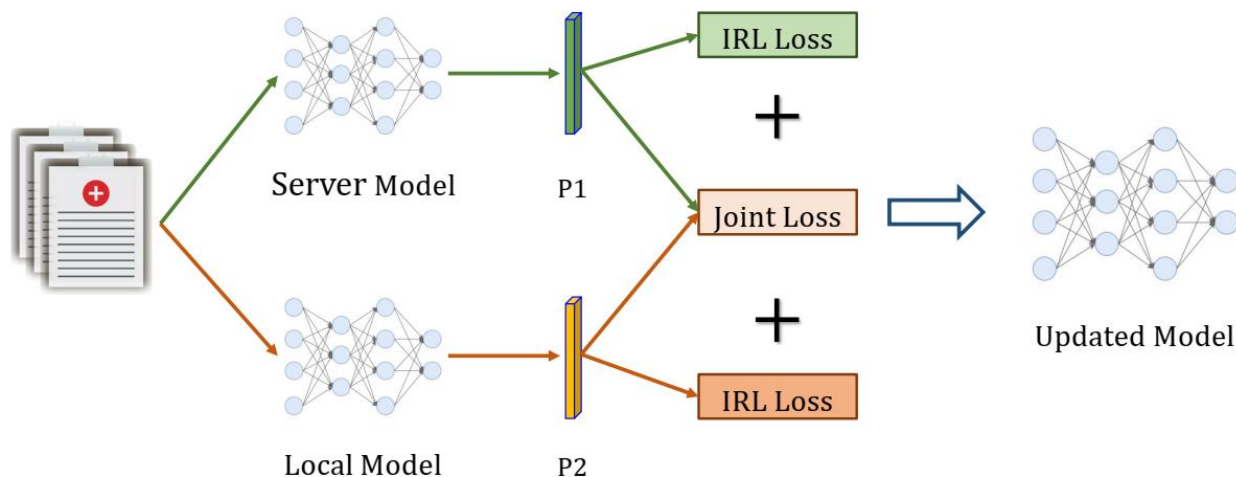
1. Zhai R, Dan C, Kolter Z, et al. Doro: Distributional and outlier robust optimization. ICML 2021.
2. Wu B, Liang Z, Han Y, et al. DRFLM: Distributionally Robust Federated Learning with Inter-client Noise via Local Mixup[J]. arXiv:2204.07742, 2022.

➤ Utilizing joint loss learning to mitigate the impact of noise.

- During classifier training, noisy data is filtered out using consistency regularization, leading to improved model robustness.

$$\mathcal{L}_{Co-Reg} = D_{KL}(P_1||P_2) + D_{KL}(P_2||P_1)$$

- Each classifier's output is compared with that of others; if similarity is found, it's deemed reliable; otherwise, it's likely noisy data.



Joint loss learning can help filter out noisy data

Evaluation



泛在无源物联网实验室
Ubiquitous Battery-free Internet of Things Laboratory



➤ Dataset: MIMIC-III

➤ Preprocess:

- Filter further to include only those admissions in which the patient was kept under ventilator support for more than 24 hours.
- Filter out admissions in which the patient is not successfully discharged from the hospital by the end of the admission

➤ The evaluation results show that FERRY improves the overall ventilation and sedation decision-making accuracy by **36.75%**

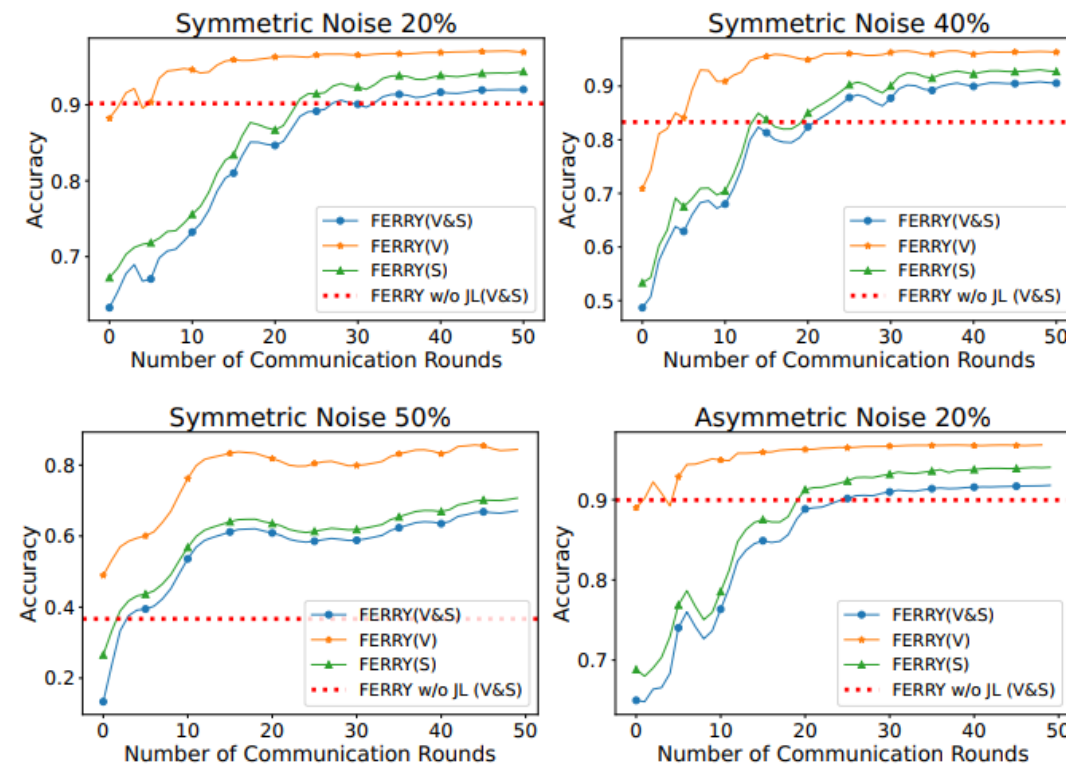


Example ventilated ICU patient.

Method	V	S	V&S	RA	Improvement(%)
CNN	86.07%	51.3%	47.52%	Not applicable	
FQI	85%	58%	55%	Not applicable	
AVRIL	96.38%	96.71%	93.43%	Not applicable	
Fed-NFQI	71.89%	16.89%	14.38%	Can't reach	Not applicable
FERRY w/o JL	96.21%	91.69%	89.16%	47	0
FERRY	96.76%	94.31%	91.75%	34	27.66

- Our method has better performance under different noise levels.
 - We create synthetic noisy label by randomly flipping the labels with various ratio.
- The results clearly demonstrate that FERRY exhibits significantly enhanced robustness against noisy labels.

Noise Level	Method					
	<i>FERRY</i> w/o JL			<i>FERRY</i>		
	V(%)	S(%)	V&S(%)	V(%)	S(%)	V&S(%)
Symmetric 20%	96.23	92.48	90.18	97.09	94.19	92.01
Symmetric 40%	90.75	85.53	83.3	96.47	93.05	90.82
Symmetric 50%	53.89	40.05	36.68	84.41	70.69	67.13
Aymmetric 20%	96.56%	92.59%	89.96%	96.83%	94.06%	91.83%



REPORT

2025



泛在无源物联网实验室
Ubiquitous Battery-free Internet of Things Laboratory



Thanks!

CSCWD 2025