DATA PROCESSING

Relational ICU database
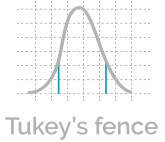
Feature selection

Time sampling
1H 4H
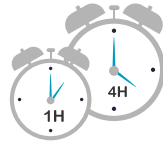
Outlier detection + Tukey's fence

Data imputation
Sample & hold + k-nearest neighbour (KNN)

Reinforcement Learning Trajectory Extraction

STATES
Clusters
Raw
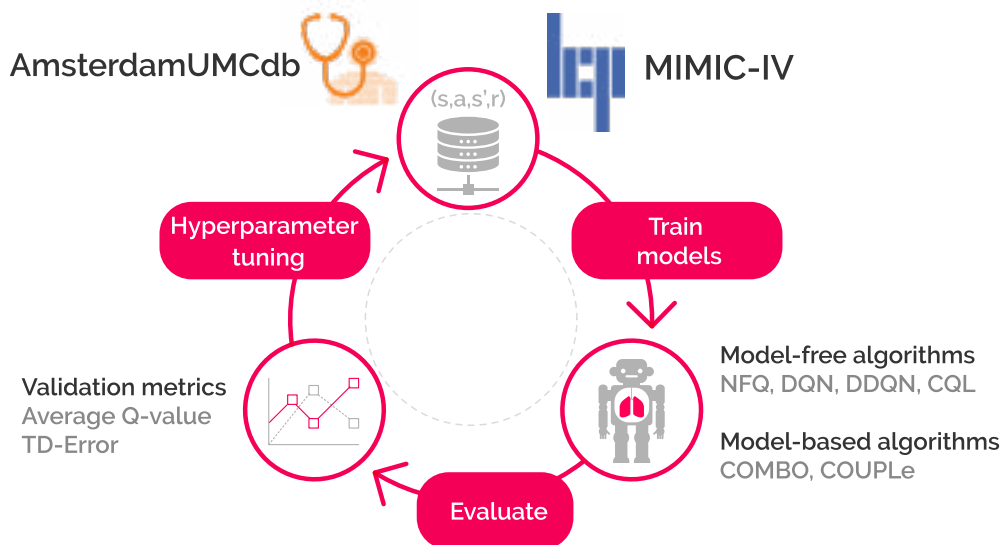Historical State Representations
Terminal Flags

ACTIONS
Discrete
Continuous

REWARDS
In-hospital and 90-day mortality
Intermediate rewards

MODEL TRAINING

AmsterdamUMCdb   **MIMIC-IV**

(s,a,s',r)

Hyperparameter tuning

Train models

Validation metrics
Average Q-value
TD-Error

Evaluate

Model-free algorithms
NFQ, DQN, DDQN, CQL

Model-based algorithms
COMBO, COUPLe

COUPLe
Conservative Optimisation using Uncertainty Penalisation for model-based q-Learning

Ensemble Dynamics Model

Predictions   Uncertainty   Dataset

$(\hat{s},\hat{a},\hat{s}',\hat{r})$ $\Psi(\hat{s},\hat{a})$ Real samples

$(\hat{s},\hat{a},\hat{s}',\hat{r} - \lambda\Psi(\hat{s},\hat{a}))$ $(\tilde{s},\tilde{a},\tilde{s}',\tilde{r})$
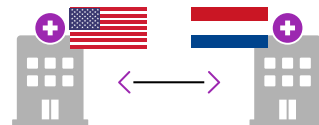
(s,a,s',r)

Conservative Q-Learning

EVALUATION
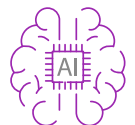
ABLATION STUDY

COMPARISON PHYSICIAN
Performance   Behaviour

OUT-OF-DISTRIBUTION

INTERPRETABILITY