Dataset Preparation and Validation for CRMAB-Based Dynamic Treatment Optimization Using MIMIC-III

Cardiac Cohort – NIT Calicut

Name	Ch. Ashish Dheer
Roll No.	M240342CS
Guide	Dr. Vidhya Kamakshi
Department	M.Tech Artificial Intelligence and Data Analytics (AIDA)
Institute	NIT Calicut
Academic Year	2024–2025
Project Title	Dynamic Treatment Optimization for Cardiac Patients using CRMAB

Abstract

This report describes the engineering of a final, causal-safe dataset for Contextual Restless Multi-Armed Bandit (CRMAB) modeling of cardiac treatment decisions using the MIMIC-III ICU database. The pipeline constructs 6-hour trajectories, aligns vitals and labs without future information leakage, removes cohort-mean artifacts, applies strictly forward-only imputation, and introduces uncertainty-aware augmentation with five draws per trajectory. The result is a validated dataset with explicit recency and missingness signals, action class weights for imbalance, and a formal feature configuration, ready for offline policy learning.

1. Background and Objectives

The objective is to build a clinically faithful state-action dataset for dynamic treatment optimization in cardiac ICU patients. The dataset must preserve temporal causality, avoid leakage, expose measurement recency, and encode actions in a reproducible form suitable for CRMAB learning.

2. Source Data

MIMIC-III v1.4 tables used include ADMISSIONS and ICUSTAYS for timelines, CHARTEVENTS for vitals, LABEVENTS for labs, INPUTEVENTS_MV and PRESCRIPTIONS for interventions, and D_ITEMS for item mappings.

3. Trajectory Construction

Trajectories are discretized at 6-hour intervals. For each admission (HADM_ID), an anchor time is established (ICU INTIME when available, otherwise ADMITTIME). For each timestep, aligned vitals and labs are associated to the nearest charted hour under analyte-specific tolerances.

4. Cardiac Cohort Selection

Admissions are retained when at least one cardiac drug was administered at any timestep. This yields a cohort aligned with diuretics, vasodilators, inotropes, and vasopressors usage.

5. Action Space

Code	Action
0	No action
1	Diuretic
2	Vasodilator
3	Inotrope
4	Vasopressor

The column `action` encodes these categories per timestep.

6. State Feature Engineering

Vitals: `hr, rr, spo2, sbp, dbp, mbp, si` (shock index), `pp` (pulse pressure). Labs: `lactate, creatinine, hemoglobin`. Recency: `ts_lactate, ts_creatinine, ts_hemoglobin`. Missingness indicators: `lactate_present, lactate_imputed` and analogues. All features are aligned to 6-hour timesteps.

6.1. Why repeated values are present

Repeated values occur naturally when a laboratory is not re-measured every timestep. The value is carried forward *causally* until a new observation arrives. This represents real ICU practice and is necessary for a Markovian state approximation between sparse measurements. Repetition caused by forward fill is clinically valid; repetition caused by global means is not.

6.2. Why 0/1 binary columns exist

Binary indicators distinguish observed versus imputed values and reveal whether a signal at a timestep is fresh or carried. These indicators improve model calibration, allow interpretability, and prevent the learner from over-trusting propagated values.

7. Missing Data Handling and Safety

Only forward fill within a patient trajectory is used. No backward fill or cohort/global mean imputation is applied. Suspected cohort-mean artifacts detected from aggressively filled versions are masked to NaN before causal processing. Measurement units are normalized and physiologic ranges enforced.

8. Uncertainty-Aware Augmentation

For timesteps where only forward-filled labs exist, five stochastic draws are generated by adding small, physiologically clipped noise estimated from observed within-patient deltas. The column `aug_id` indexes these draws (1–5). This approximates multiple imputation and improves robustness without violating causality.

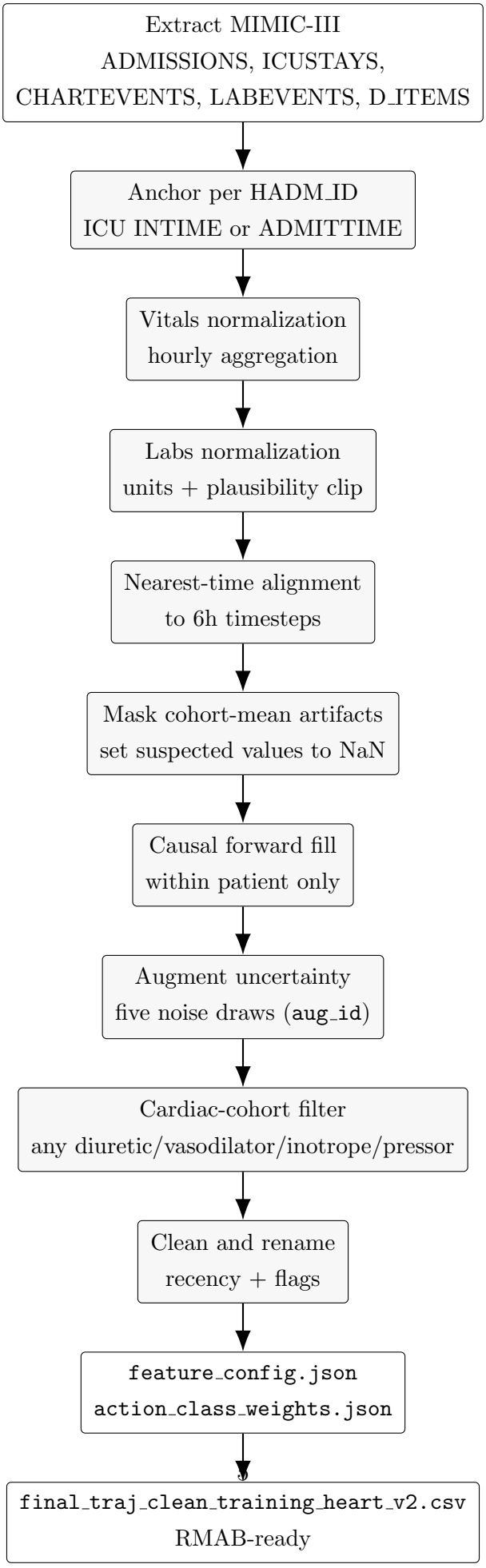
9. Feature Configuration and Class Weights

The file `feature_config.json` enumerates the state features and their order for reproducibility across training and evaluation. The file `action_class_weights.json` stores inverse-frequency class weights to mitigate action imbalance, ensuring rare but critical actions receive adequate learning signal.

10. Final Dataset Summary

Final CSV	`final_traj_clean_training_heart_v2.csv`
Patients	4,927
Rows	197,080
Timestep size	6 hours
Augmentations	5 (`aug_id`=1–5)
Causality and leakage	Forward-only, no backfill, no cohort means
Readiness	RMAB-ready

11. Preprocessing Flowchart



12. RMAB Formulation

Let $s_t \in \mathcal{S}$ denote the patient state at timestep t, $a_t \in \mathcal{A} = \{0, 1, 2, 3, 4\}$ the action, and $r_t = R(s_t, a_t)$ the reward. A stationary policy $\pi(a \mid s)$ maps states to action probabilities. The objective is to find a policy that maximizes the expected discounted return

$$\max_{\pi} \mathbb{E}\left[\sum_{t=0}^{T} \gamma^t r_t \,\middle|\, \pi\right], \quad \gamma \in (0, 1].$$

In CRMAB settings, multiple patients evolve in parallel with partially observed, restless dynamics. The state vector includes vitals, labs, recency features, and missingness indicators to preserve causal signal and measurement freshness.

13. Reward Design Sketch

A simple, clinically plausible reward can combine physiological stability and action cost. One example is

$$R(s_t, a_t) = \alpha \cdot f_{\text{hemodynamics}}(s_t) + \beta \cdot f_{\text{metabolic}}(s_t) + c(a_t),$$

where $f_{\text{hemodynamics}}$ rewards lower shock index and adequate blood pressure, $f_{\text{metabolic}}$ rewards lower lactate, and $c(a_t)$ penalizes interventions with increasing cost for inotropes and vasopressors. The reward can be refined with outcome-aligned terms as needed.

14. Why This Dataset Is Final

It preserves temporal causality with forward fill only, eliminates cohort-mean artifacts, exposes recency and missingness via explicit features, normalizes units within physiologic bounds, addresses class imbalance with action weights, and uses a formal feature configuration for reproducibility. Repeated values reflect real ICU measurement frequency rather than leakage. Binary 0/1 flags enable interpretability and robust modeling. The dataset is therefore approved as the final state-action basis for CRMAB policy learning.

15. Limitations and Future Work

Outcome-linked reward refinement and explicit transition modeling can further improve policies. Additional labs or static covariates may be incorporated by extending the feature configuration. External validation on other cohorts is recommended.