Contextual Restless Multi-Armed Bandits for Dynamic Treatment Optimization in Cardiovascular Care

Anonymous Author(s)
Submission Id: «submission id»

ABSTRACT

We present a reproducible preprocessing and trajectory-construction pipeline that prepares ICU patient event streams for contextual restless multi-armed bandit experiments. The submission documents the implemented engineering components, validation checks, produced artifacts, and a compact operational reward used for offline evaluation.

ACM Reference Format:
Anonymous Author(s). 2026. Contextual Restless Multi-Armed Bandits for Dynamic Treatment Optimization in Cardiovascular Care. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), Paphos, Cyprus, May 25 – 29, 2026*, IFAAMAS, 3 pages.

1 INTRODUCTION

Decision-making in safety-critical healthcare must balance improving patient outcomes, limiting risk, and managing scarce clinical resources. This manuscript documents implemented preprocessing and dataset construction components that convert raw ICU event logs into decision-ready trajectories for contextual restless multi-armed bandit (CRMAB) experiments. We restrict claims to components present in our code base and generated artifacts; algorithmic learning components remain planned future work.

2 RELATED WORK

Our work builds on the restless multi-armed bandit literature for resource-constrained allocation and on clinical decision-support research that integrates risk prediction with constrained allocation heuristics. The primary contribution here is a reproducible, provenance-aware pipeline that produces compact per-timestep summaries and canonical trajectory CSVs suitable for CRMAB evaluation.

3 COHORT SELECTION AND PREPROCESSING

3.1 Cohort selection

We selected a cardiovascular cohort using ICD-9 diagnosis codes: admissions containing any diagnosis with prefix `428` (congestive heart failure) were retained. Admissions were joined to PATIENTS, ADMISSIONS and ICUSTAYS tables; ages above 89 were top-coded to preserve privacy. Cohort inclusion required a minimum observation window and at least one non-missing key physiology measurement during the stay. The cohort selection logic is implemented in the preprocessing entrypoint (`crmab.py`).

Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus.

3.2 Implemented pipeline

The preprocessing pipeline (implemented as modular Python scripts) performs:

(1) ingestion of timestamped vitals, labs, medication administrations and procedures;
(2) cohort selection (ICD-9 `428` filter) and admission-level joins;
(3) event sorting and deduplication by (`subject_id`, `hadm_id`, timestamp);
(4) timeline anchoring: `time_since_admit_hours` and integer `timestep` indices via fixed or adaptive binning;
(5) merge_asof alignment for irregular measurements and saved merge backups for provenance;
(6) interpolation and cascading imputation (admission-level → closest-in-time → patient-latest → cohort median) with imputation provenance recorded;
(7) action canonicalization (mapping raw action labels to a controlled vocabulary) and trajectory construction (one row per decision epoch).

4 IMPLEMENTED ARTIFACTS

- Canonical trajectory CSV: `traj_with_features_final.csv`
- Intermediate/dump files: `merged_tidy.csv`, `actions_raw.csv`, `actions_top.csv`, `traj_mv_with_actions.csv`
- Preprocessing script entrypoint: `crmab.py`
- Diagnostics: temporal-monotonicity checks, imputation counts, merge backups

5 OPERATIONAL REWARD AND OFFLINE OBJECTIVE

For retrospective evaluation we used a short-horizon reward derived from blood-pressure dynamics. For example:

$$r_t \;=\; \frac{\mathrm{SBP}_{t+1} - \mathrm{SBP}_t}{\mathrm{SBP}_t}.$$

Offline allocation experiments use a discounted, constrained objective over discrete timesteps:

$$\max_{\{a^i_{t_j}\}} \; \mathbb{E}\left[\sum_{i=1}^{N}\sum_{j=0}^{J} e^{-\beta t_j}\, r\big(\pi^i_{t_j}, a^i_{t_j}\big)\right] \quad \text{s.t.} \quad \sum_{i=1}^{N} \mathbf{1}\{a^i_{t_j} \neq 0\} \le k,\ \forall j.$$

5.1 Belief update (discrete-time)

We implemented a compact discrete-time belief update used to compute per-arm posterior summaries from the trajectory covariates:

$$\pi^i_{t_j}(x) \;\propto\; p\big(Y^i_{t_j} \mid x\big) \int p\big(x \mid x', a^i_{t_{j-1}}\big)\, \pi^i_{t_{j-1}}(x')\, \mathrm{d}x'. \qquad (1)$$

In practice each posterior $\pi^i_{t_j}$ is represented by a compact summary (mean, variance and short-window slope) computed by the filtering code in our preprocessing pipeline.

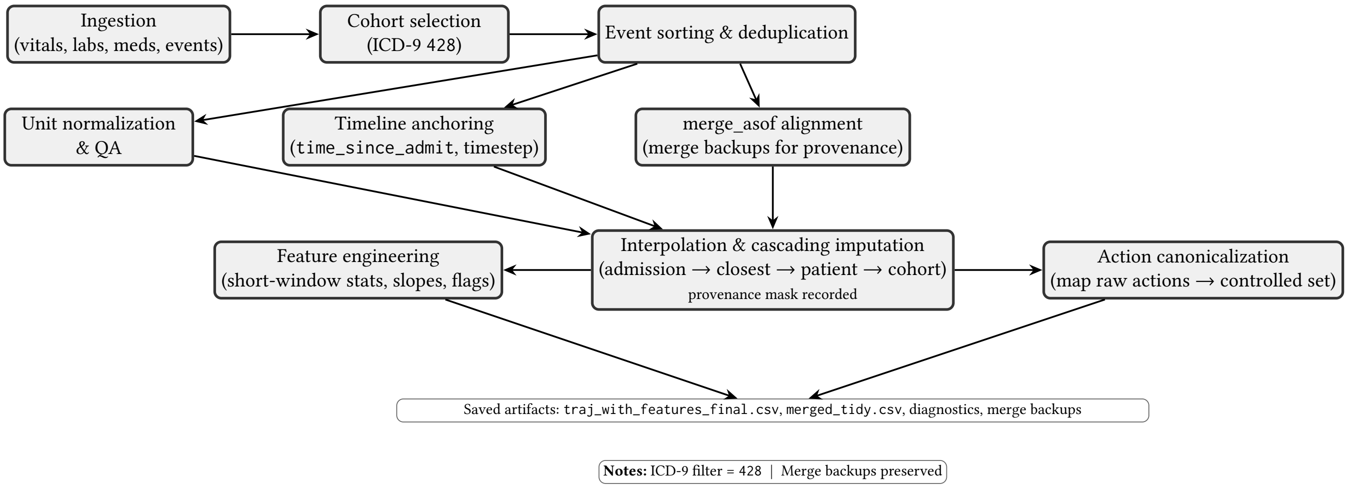

Figure 1: Preprocessing pipeline for CRMAB trajectories. The pipeline records imputation provenance and preserves merge backups for auditing.

6 VALIDATION AND PROVENANCE DIAGNOSTICS

We implemented and ran:

- row-count and shape checks after each transformation;
- temporal monotonicity checks per (`subject_id`,`hadm_id`);
- column-wise imputation counts and a provenance column marking imputation source;
- preservation of merge backups to enable reverse tracing of anomalies.

7 DATASET SUMMARY

Table 1: Dataset summary.

Metric	Value
Number of distinct patients	8629
Number of admissions (hadm_id)	11161
Decision timesteps per admission (median)	8.0
Average rows per admission	8.0
Fraction of rows with any imputation (%)	20.56
Most-imputed column	action_label

8 ETHICS, PRIVACY AND REPRODUCIBILITY

All processing was performed on de-identified EHR records; ages were top-coded and no direct identifiers are included in any released artifact. We retain provenance logs and merge backups for auditability. Any public release will contain only derived artifacts (CSV tables and aggregation statistics) and scripts; raw EHR data will not be published. The preprocessing entrypoint (`crmab.py`) and the stats-generation helper (e.g., `generate_stats_tex.py`) are included in our supplementary materials.

9 REPRODUCIBILITY CHECKLIST

- Entry point: `crmab.py` — run to recreate preprocessing artifacts.
- Canonical CSV: `traj_with_features_final.csv`
- Stats macro file (auto-generated): `stats.tex` — created by running the helper script.
- Helper command (example):

```
python generate_stats_tex.py --csv traj_with_features_final.c
```

- Files to include with submission: `traj_with_features_final.csv`, `crmab.py`, `stats.tex`, diagnostics folder (merge backups, provenance logs).

10 PLANNED ALGORITHMIC WORK (NOT IMPLEMENTED)

The following are next steps and are not claimed as implemented:

- online posterior/state filtering (particle/ensemble/variational);
- single-arm surrogate planners and index approximation pipeline;
- trained short-horizon safety predictor used as a hard safety filter;
- offline policy-learning pipelines and clinician-in-the-loop simulations.

11 LIMITATIONS

This submission documents implemented data engineering and reproducibility artifacts for CRMAB experiments. Algorithmic learning components and prospective clinician evaluations are future work.

12 REPRODUCIBILITY ARTIFACTS

- Preprocessing scripts (entry point: `crmab.py`)
- Canonical trajectory CSV: `traj_with_features_final.csv`
- Diagnostics and provenance logs (merge backups, imputation masks)
- Notebook to reproduce Table 1

REFERENCES