Brief Report: Predicting Weekly Hospital Crisis Periods

Team:
Antoine Moussa
Dina Abi Akar
Youssef Al Hajj Youness
Zenab Bassam
Rayna Mhanna

October 28, 2025

Project Overview

We frame the task as a binary classification problem: for each calendar week, predict whether conditions constitute a *crisis* (label y=1) or not (y=0). A week is labeled crisis if operational thresholds are exceeded (e.g., very high occupancy 98% and/or a surge in refused patients > 40). The goal is to provide early signals that support staffing and capacity decisions while keeping the model simple, auditable, and robust.

Data and Features

The working dataset aggregates historical hospital signals at a weekly cadence. Core predictors include: overall occupancy/load measures, counts of patient refusals, service/department mix, and temporal indicators (week-of-year, holidays). Records are cleaned (type casting, missing handling, outlier capping when appropriate) and split into train/test with stratification by the target to preserve class balance.

Modeling Approach

We benchmarked several supervised learners (logistic regression, tree-based gradient boosting) and selected the model-family that balanced discrimination performance and interpretation. Threshold selection is based on precision–recall trade-offs aligned to operational costs (missed-crisis vs. false-alarm). We compute standard metrics on a held-out set: AUROC, AUPRC, precision, recall, F1, and confusion matrix, and verify stability with cross-validation.

Explainability and Diagnostics

To ensure decision transparency, we use SHAP values to attribute each prediction to input features. Typical high-impact drivers are elevated occupancy and spikes in refusals; temporal factors modulate risk seasonally. Calibration is checked (reliability curve) to justify probability-based thresholds. Error analysis focuses on borderline weeks to refine rules and thresholds.

Operationalization

Predictions can be generated weekly from the latest data snapshot. We export probabilities and a binary flag based on the chosen threshold, plus the top contributing features per prediction

for review. A lightweight dashboard can show recent risk, feature attributions, and alerts when risk surpasses the agreed threshold.

Limitations and Next Steps

Data representativeness (e.g., policy changes, coding changes) may affect generalization. Future work: incorporate exogenous signals (seasonality, public holidays), refine the crisis definition in consultation with clinicians/operations, and evaluate lead-time forecasting (predicting crisis k weeks ahead).

Reproducibility. Experiments are run with fixed random seeds and stratified splits; code notebooks document preprocessing, training, and evaluation.