
Predicting ED Stay and Disposition with Interpretable ML



Emergency departments (EDs) face increasing challenges related to patient flow, overcrowding and resource allocation. Predicting key outcomes such as patient length of stay (LOS) and disposition decision (DD) can significantly improve operational efficiency. However, existing machine learning (ML) models often lack transparency and transferability, limiting their broader applicability. A recent study published in BMJ Health & Care Informatics aimed to develop interpretable ML models capable of predicting LOS and DD at various time points post-triage, while establishing a transparent data analysis framework that other institutions can adopt and customise.

A Transparent Framework for Predictive Modelling

The research focused on a quaternary hospital in Melbourne, Australia, analysing over 297,000 ED visits between mid-2019 and the end of 2022. After applying exclusion criteria to remove visits involving deaths, minors, patients who left prematurely and COVID-19 testing-only cases, the data was divided into modelling and testing sets. Predictions targeted two main outcomes—LOS and DD—at three time points (10, 60 and 120 minutes after triage), resulting in 12 individual models. The LOS was predicted as both a binary outcome (above or below four hours) and a ternary outcome (≤ 4 , 4–24, >24 hours), while the DD covered discharge, short stay or inpatient admission/transfer.

Feature processing involved categorising and encoding both categorical and numeric values, with novel methods used to handle missing values and U-shaped correlations. Notably, instead of using one-hot encoding, categorical variables were target encoded to reduce dimensionality. For numeric features, values were divided into intervals based on their predictive relationship with the outcome. This process enabled the use of generalised linear models such as lasso regression while maintaining model interpretability.

Interpretable Model Development and Validation

The final models relied on a streamlined selection of 21 features, chosen through lasso regression to mitigate multicollinearity. Eight ML algorithms were assessed—logistic, ridge, lasso, elastic net, decision tree, random forest, xgboost and an ensemble model. Despite slight performance differences among them, the lasso models were selected due to their balance of accuracy and interpretability.

Must Read: [Reducing Hospital Length of Stay Through Innovation](#)

Testing results showed robust predictive performance. Binary LOS prediction models reached area under the curve (AUC) values of 0.862, 0.868 and 0.878 at 10, 60 and 120 minutes respectively. Binary DD prediction models achieved AUCs of 0.839, 0.851 and 0.863. For ternary predictions, LOS models had accuracies ranging from 60.2% to 61.9%, and DD models from 61.5% to 63.4%. Additionally, cross-validation helped define optimal cut-off points, further enhancing predictive accuracy.

Model predictions proved useful for real-world decision support. For example, one patient case showed a predicted LOS probability of 0.973 at 120 minutes, far exceeding the 0.429 threshold. This matched the actual outcome—an extended hospital stay of over 40 days—demonstrating the model's alignment with clinical reality. Key features like 'Order count', 'Age' and 'Average waiting time' emerged as strong predictors of LOS and DD, with many overlapping between both outcomes. This reflects the interdependence of LOS and patient disposition within clinical workflows.

Impact and Applicability in Clinical Settings

The strength of this study lies not only in model accuracy but in the clarity of its data processing framework. Each modelling step—from data pre-
© For personal and private use only. Reproduction must be permitted by the copyright holder. Email to copyright@mindbyte.eu.

processing to encoding, feature selection and validation—was explicitly documented. This allows healthcare institutions with different datasets to adopt the framework and tailor their own interpretable models. It also supports future enhancements like drift adaptation, addressing the problem of model performance decay over time due to changes in hospital operations or patient populations.

Importantly, interpretability ensures that predictions can be understood and trusted by clinicians. Rather than relying on opaque black-box systems, the lasso-based models allow clinicians to examine which features influenced the predictions. This supports informed decision-making, promotes clinician engagement and facilitates communication with patients and staff. Features like 'Average waiting time' significantly influenced LOS but not DD, offering nuanced insights that reflect real ED dynamics. Such differentiation helps target interventions more effectively, such as improving throughput or preparing for patient admission.

The framework's adaptability also counters the recognised limitations of global models, which tend to underperform in site-specific contexts. By enabling localised model development, the approach enhances both accuracy and clinical relevance. This is particularly critical in high-stakes environments like EDs, where timely and accurate predictions can have direct implications for patient safety and care delivery.

The study successfully developed 12 interpretable ML models for predicting LOS and DD in the ED at different time points post-triage. Through meticulous data handling and model selection, the research established a transparent and adaptable framework for clinical prediction modelling. The lasso-based models not only performed well but offered clarity and trustworthiness for clinical use. By empowering healthcare institutions to develop their own models using this methodology, the study offers a scalable solution to improve decision-making and resource management in emergency care.

Source: [BMJ Health & Care Informatics](#)

Image Credit: [iStock](#)

Published on : Sun, 6 Jul 2025