Predicting Emergency Department Disposition Using Statistical Learning

Statistical Learning for ANLY Final Project

Wendy Hu, Christy Hsu

I. Abstract

Emergency department (ED) disposition—the decision of whether a patient is admitted, discharged, transferred, leaves voluntarily, or returns—plays a critical role in patient care and hospital resource management. This study applies machine learning techniques to predict ED disposition outcomes using patient-level data from the 2022 National Hospital Ambulatory Medical Care Survey (NHAMCS), comprising over 16,000 visits and 900 variables. We performed extensive data cleaning, consolidation of 11 original disposition labels into five clinically meaningful categories, and applied Lasso-based feature selection.

Four classification models were evaluated: Logistic Regression with Lasso, Support Vector Machine (SVM), Random Forest, and XGBoost. XGBoost achieved the highest predictive performance overall, while Random Forest provided a strong balance of accuracy and interpretability. Notably, the multinomial Logistic Regression model with Lasso regularization yielded the smallest gap between training and test performance, demonstrating strong generalizability. It also enabled deeper interpretability by identifying class-specific predictors across clinical, socioeconomic, and procedural categories.

Our findings demonstrate that ED disposition is primarily driven by real-time clinical factors—particularly patient vitals, procedures, and medication use—rather than demographic attributes. This supports the development of fair and effective decision-support tools to improve patient flow and care efficiency in emergency settings.

II. Introduction

Disposition — the decision regarding a patient's next step after evaluation in the emergency department (ED) — plays a vital role in patient care and hospital operations. A correct disposition ensures that patients receive appropriate treatment pathways, whether it be admission, discharge, observation, or transfer. Beyond individual outcomes, disposition decisions have a profound impact on the overall efficiency of the hospital system.

Delayed or inaccurate dispositions can lead to ED overcrowding, bed shortages, prolonged wait times, and strained resources, ultimately affecting the quality of care for all patients. Because of the high stakes and complexity, disposition decisions require balancing clinical judgment, patient-specific factors, and institutional constraints.

Improving disposition prediction is essential for optimizing patient flow, enhancing healthcare delivery, and supporting the resilience of emergency services. In this study, we leverage real-world emergency room data to identify key predictors of patient disposition, applying feature selection and machine learning models to uncover patterns that could inform better, faster, and more reliable decision-making in emergency care.

III. Literature Review

Research on implementing or exploring the potential of machine learning tools in Emergency Department (ED) settings has primarily focused on three major areas of interest: First, triage and acuity scoring, where machine learning models are used to evaluate existing severity indexing, provide assistance to triage ranking system, or propose new scoring measures[1].Second, the efficiency of ED resource usage[2], studies apply machine learning algorithms to investigate patient length of visit, predict ED patient flow and usage. Third, visit outcomes, where most research targets the prediction of admission disposition. Hong, Haimovich, Taylor's 2018 paper framed their ML task as a binary classification between admission and discharge, offering us guidance for grouping complex patient records and hospital characteristics into broader categories[3]. A recent research on ED disposition prediction is the 2025 paper by Savaş, Cingiz and Ibiş which set forth for finding multi-class classifiers to discriminate between referral, discharge, hospitalization, and mortality; we reference their choice of learning algorithms as it aligns with the design of our response variable[4].

In our project, we made a particular design for our response variable, expanding previous researchers' interests on admission, mortality and referral outcomes. Namely, we distinguish special dispositions, such as cases who returned to ED within 72 hours, and cases who left the ED early without completing the formal discharge process, as classes that stand alone. This decision was inspired by the reading of the 2017 report from DHHS (U.S. Department of Health and Human Services), which offers an accessible and in-depth discussion on the issue of "discharge failure" in EDs.[5]

IV. Data Collection

1. Data Source

We employed the 2022 National Hospital Ambulatory Medical Care Survey (NHAMCS) ED dataset, which provides rich information across demographics, triage evaluations, clinical interventions, hospital characteristics, and outcomes[6].

To supplement our data, we used the ICD-10-CM (International Classification of Diseases, 10th Revision, Clinical Modification) dataset, which helps us decode the provider's diagnoses, causes of injury, and the reason for visit feature.

2. Data Preprocessing

Through literature review, we selected 188 clinically relevant features from an initial pool of 913. These features were categorized into domains such as Demographics, Triage, Injury, Hospital Usage, Hospital features, Chief Complaint, Diagnosis, and Procedures, Chronic Disease. During preprocessing, we identified that the original dataset contained 11 distinct patient disposition categories. To facilitate more robust model training and avoid sparsity issues, we consolidated these into five clinically meaningful groups:

- **Admit** - Patients who were admitted to the hospital for further treatment following their ED visit. This includes direct hospital admissions and observation cases that resulted in admission or discharge. The total count of the admitted patients in the dataset is 2302.
- **Discharge** - Patients who were discharged home after receiving emergency care, without requiring follow-up hospitalization. The total count of the discharged patients in the dataset is 1202.

- **Left Early** - Include patients who left the emergency department before treatment completed, patients who left against medical advice and patients who left voluntarily without being seen. The total count of the patients who left early is 748.
- **Returned** - Patients who initially were discharged but returned to the ED within a short window, typically within 72 hours. The total count of this type of patients is 2053.
- **Transfer** - Patients who were transferred from the current emergency department to another medical facility for specialized care or resource availability reasons. The total count of the patients in this category is 452.

This regrouping aimed to strike a balance between clinical relevance and modeling feasibility, addressing class imbalance and reducing label noise to enhance predictive performance.

3. Data Cleaning

Given the size and complexity of the dataset, data cleaning and preprocessing were essential steps before modeling. For categorical variables, missing entries and ambiguous responses (such as "Blank" or "Not Applicable") were recoded as "Unknown" or "Missing" to retain these cases without introducing bias. Categorical variables were subsequently transformed into machine-readable format using one-hot encoding, allowing our models to properly interpret categorical information without imposing ordinal assumptions.

For continuous variables, we applied random imputation, replacing missing values by randomly sampling from the observed, non-missing entries of the same variable. This approach preserves the distribution of each feature while avoiding the loss of information that would result from casewise deletion or mean imputation.

Following these preprocessing steps, we generated cleaned and encoded datasets ready for machine learning modeling, ensuring consistent feature formats and minimizing the potential for missing data biases.

V. Methods

To predict emergency department (ED) patient disposition, we applied four supervised learning models: Logistic Regression with Lasso regularization, Support Vector Machine (SVM), Random Forest, and XGBoost. Each model was selected based on its theoretical foundation, relevance to structured healthcare data, and potential to balance performance with interpretability.

1. Model Techniques and Applications

Logistic Regression: Logistic Regression with Lasso (L1) regularization was chosen as our baseline model for its simplicity and high interpretability. This method estimates the probability of a case belonging to a particular class by modeling the log-odds as a linear function of the input variables. To mitigate overfitting and improve feature selection, we applied Lasso regularization, which penalizes the absolute value of the coefficients and forces less informative ones to zero. While logistic regression assumes linearity in feature effects and independence between predictors, its transparency allowed us to directly quantify how specific variables influenced the likelihood of each disposition outcome.

Support Vector Machine: To capture potential non-linear relationships and subtle boundaries between disposition classes, we employed a Support Vector Machine (SVM) with a radial basis function (RBF) kernel. SVM constructs a maximum-margin hyperplane to separate classes in a high-dimensional space, and the RBF kernel enables it to handle non-linear separability. This was particularly useful in distinguishing classes with overlapping clinical profiles. However, SVMs require extensive tuning and are sensitive to class imbalance, which impacts performance in our multi-class setting.

Random Forest: We also utilized Random Forest, an ensemble learning method that constructs a large number of decision trees and aggregates their predictions. Each tree is trained on a bootstrap sample with a random subset of features considered at each split. This technique naturally models variable interactions, reduces overfitting, and outputs feature importance measures. In our study, Random Forest highlighted the predictive strength of variables, while confirming the relatively low contribution of demographic features. Although less interpretable than logistic regression, it provided more robustness and flexibility in modeling complex patient profiles.

XGBoost: Finally, we implemented XGBoost (Extreme Gradient Boosting), a highly optimized gradient boosting framework that sequentially builds decision trees to minimize prediction errors from previous iterations. XGBoost supports both L1 and L2 regularization, handles missing values internally, and is known for its efficiency and scalability. In our experiments, XGBoost achieved the highest overall performance across multiple evaluation metrics (Accuracy, F1, AUC), especially in handling underrepresented disposition classes like "Returned" or "Transfer." It effectively captured higher-order feature interactions, such as combinations of chronic disease status, medications administered, and triage acuity, which were strongly associated with admission likelihood.

2. Model Training Procedure, Cross-Validation, and Hyperparameter Tuning

All models were trained and evaluated using stratified 5-fold cross-validation, preserving class distributions within each fold. This ensured that performance estimates remained stable despite imbalanced class sizes.

For Logistic Regression with Lasso regularization, we employed a regularization path approach to identify the optimal penalty parameter (λ). A grid of λ values was tested using internal cross-validation within each training fold. We selected the λ value that minimized the mean cross-validated log loss, thereby balancing model complexity with predictive accuracy. This procedure allowed us to retain only the most informative features and shrink the coefficients of irrelevant variables to zero.

The Support Vector Machine (SVM) model used a radial basis function (RBF) kernel to capture non-linear decision boundaries between disposition categories. Hyperparameters including the cost parameter (C) and kernel width (gamma) were tuned via grid search, testing combinations such as C = 0.1, 1, 10 and gamma = 0.01, 0.1, 1. For each configuration, 5-fold cross-validation was used to evaluate performance based on macro-averaged F1 score, which was chosen due to its ability to reflect performance across all disposition classes. The final model was selected based on the configuration that yielded the highest cross-validated F1 score.

The Random Forest model was tuned by varying the number of trees (`n_estimators`) and the maximum number of features considered at each split (`max_features`). We tested configurations with 100, 250, and 500 trees, and `max_features` set to either the square root or the logarithm of the total number of predictors. Cross-validation was used to identify the combination that maximized recall and AUC across folds. The final model included 500 trees and used the square root of the feature set size at each split, which provided a good trade-off between performance and overfitting control.

XGBoost underwent a more detailed tuning process due to its sensitivity to hyperparameter choices. We first performed coarse tuning on key parameters including the learning rate (`eta`), tree depth (`max_depth`), and subsampling rate (`subsample`). Once approximate ranges were established (e.g., learning rate = 0.1–0.3, `max_depth` = 4–8), we used a randomized grid search combined with 5-fold cross-validation to refine these values. Additional regularization parameters, such as `gamma` (minimum loss reduction) and `lambda` (L2 penalty), were adjusted to prevent overfitting. Early stopping rounds were also applied within each fold to halt training when validation performance plateaued, typically after 50–100 boosting rounds. The final XGBoost model incorporated a learning rate of 0.2, `max_depth` of 6, and 300 estimators, with early stopping enabled.

All modeling was performed using R, leveraging the `glmnet`, `e1071`, `randomForest`, and `xgboost` packages. Predictions were made on held-out test sets after cross-validation, and performance metrics were calculated using multi-class variants of Accuracy, Precision, Recall, F1 Score, and AUC.

By combining both interpretable models and state-of-the-art ensemble methods, our modeling framework allowed us to both understand the factors that drive ED disposition decisions and achieve strong predictive performance across diverse patient cases.

VI. Results

1. Model Performance Evaluation

To assess model effectiveness in predicting emergency department (ED) disposition outcomes, we trained and evaluated four classification models: Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), and XGBoost. The performance of each model was evaluated using five standard metrics: Accuracy, Precision, Recall, F1 Score, and Area Under the ROC Curve (AUC).

Model	Accuracy	Precision	Recall	F1 Score	AUC
LR	0.7242	0.6809	0.6037	0.6223	0.8726
SVM	0.5085	0.5990	0.5764	0.5225	0.8981
RF	0.7146	0.6520	0.5956	0.6087	0.8932
XGBoost	**0.7376**	0.6790	**0.6176**	**0.6363**	**0.8893**

TABLE I: Comparison of model performance metrics on the test set.

Among the four model performances, **XGBoost** achieved the highest overall accuracy (73.76%) and F1 score (0.6363), indicating strong predictive capability, particularly in balancing performance across all disposition classes. However, its relatively higher performance variance between training and test sets raised concerns about potential overfitting.

Support Vector Machine (SVM), while showing a high AUC (0.8981), demonstrated lower accuracy (50.85%) and F1 score (0.5225), suggesting limited generalizability, especially given the class imbalance and complexity of ED disposition patterns.

Random Forest offered a favorable balance between predictive performance and robustness, achieving a competitive accuracy of 71.46% and a strong AUC of 0.8932, while maintaining consistent recall and F1 scores across all classes.

Logistic Regression, although not the top performer in terms of predictive metrics, exhibited the most stable behavior, with minimal discrepancy between training and validation scores—making it a dependable and interpretable baseline model.

To further assess model discrimination, we plotted the Receiver Operating Characteristic (ROC) curves for all models. Fig. 1 illustrates that all models perform well above the chance line (diagonal), with steep initial rises in true positive rate, indicating strong sensitivity. XGBoost and SVM slightly outperform the others in terms of AUC, yet Random Forest's curve remains comparably high across all thresholds. Given these findings, we selected Random Forest for its predictive strength and Logistic Regression for its transparency and consistency to further examine feature importance and explore the variables most influential in predicting ED disposition outcomes.

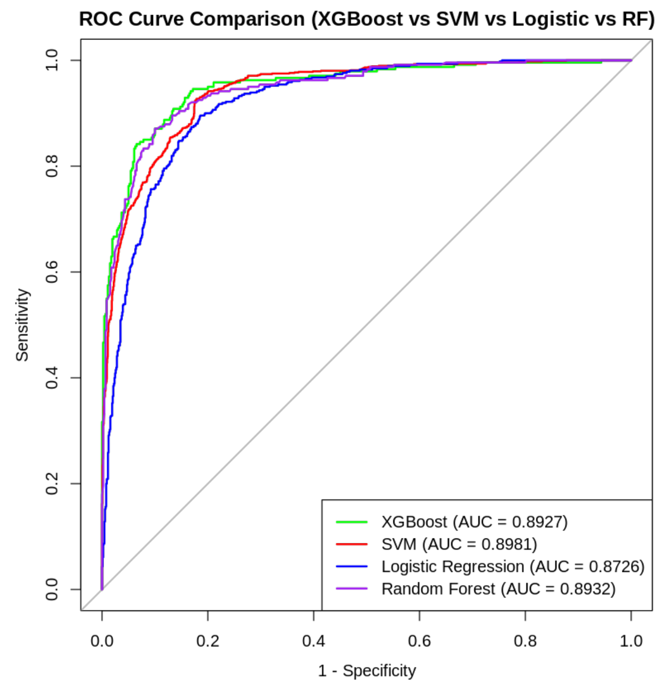


Fig. 1: ROC Curve Comparison (XGBoost vs SVM vs Logistic vs RF)

2. Random Forest

Confusion Matrix: To better understand the performance of the Random Forest model beyond overall accuracy, we examined the confusion matrix, which provides a detailed breakdown of predicted versus actual classes across all disposition categories. This approach allows us to assess not only how often the model correctly predicts each class, but also where systematic misclassifications occur. In multi-class classification problems like ours—where the categories are clinically nuanced and sometimes overlapping—confusion matrices are especially valuable for uncovering class-specific weaknesses. By analyzing which categories are most frequently confused, we gain insights into the limitations of the model, as well as potential real-world factors (such as hospital resource constraints or ambiguous clinical presentations) that contribute to predictive uncertainty.

Fig 2 reveals that the two most frequent disposition categories—"Admit" and "Returned"—are generally well predicted. While misclassifications are present across other categories, they offer valuable insights into the clinical and operational nuances behind ED disposition.

The most common misclassification involves cases that were predicted as "Admit" but were actually "Transfer." This confusion may stem from the similarity in clinical presentation: both groups typically involve patients with serious or complex conditions requiring continued care. The key distinction likely lies in hospital resource availability. For instance, patients requiring specialized care—such as psychiatric services or long-term skilled nursing—may be transferred to external facilities if the original ED lacks the necessary infrastructure. Thus, the decision between admission and transfer may not be purely clinical but also operational, depending on bed availability, staffing, or specialty services.

The second most frequent misclassification occurred when the model predicted "Returned" but the true label was "Discharge." This highlights the challenge of anticipating which discharged patients will come back within a short timeframe. Patients may return due to unresolved symptoms, progression of illness, inadequate treatment, or even misunderstanding of discharge instructions. While some of these returns may reflect the limitations of initial diagnosis or treatment decisions, others could be tied to patient-level factors such as lack of access to follow-up care, non-compliance with medication, or social determinants

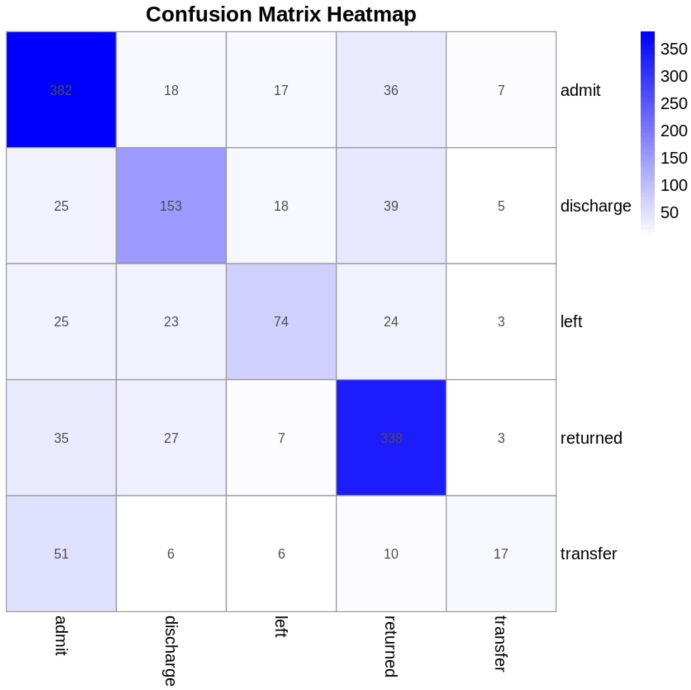


Fig. 2: Confusion Matrix Heatmap

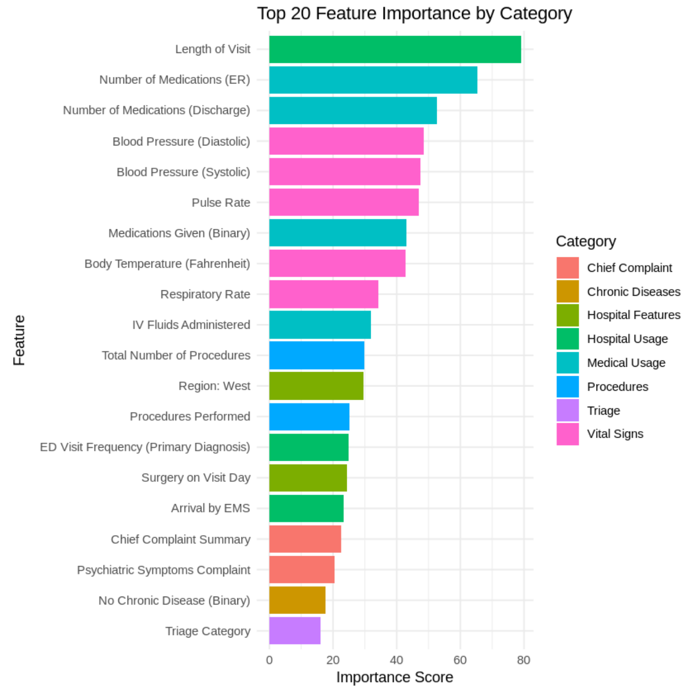


Fig. 3: Top 20 Feature Importance by Category

of health (e.g., homelessness, caregiving burdens). We assume that some “Returned” patients were not clinically distinguishable from those who were discharged, further complicating the model’s ability to differentiate these two outcomes solely from visit-level features.

This analysis underscores the fact that certain disposition outcomes—particularly “Transfer” and “Returned”—may be driven by contextual factors not captured in the structured dataset, such as hospital capacity, real-time staffing, and patient follow-up support, which should be considered in future model iterations.

Top 20 feature selection: To further interpret the Random Forest model, we examined the top 20 most important features contributing to disposition prediction, as shown in the feature importance plot. Fig 3 clearly highlights that variables reflecting a patient’s clinical condition—including vital signs (e.g., blood pressure, pulse rate, body temperature), medication usage, and procedural indicators—play the most prominent role in determining disposition outcomes. Notably, Length of Visit and Number of Medications (both in the ER and at discharge) emerged as the top predictors, underscoring the relevance of treatment intensity and visit complexity in clinical decision-making.

To aid interpretation, features were color-coded by category. This color encoding reveals that vital signs, medical usage, and procedures dominate the top ranks, whereas demographic and administrative variables such as Region or Arrival by EMS are less prominent. These findings reinforce the conclusion that disposition decisions are primarily driven by real-time patient condition and clinical management rather than by static background characteristics.

To further explore how specific clinical features influence disposition predictions, we examined Partial Dependence Plots (PDPs) for two of the most important variables identified by the Random Forest model: Length of Visit (LOV) and Pulse Rate (PULSE). These plots visualize the marginal effect of each variable on the model’s predicted probability for each disposition category, while averaging out the effects of all other features.

Fig 4 reveals a strong and interpretable relationship with patient outcomes. As the duration of the ED stay increases, the predicted probability of admission and transfer rises sharply, particularly after a certain threshold (e.g., several hours). This pattern aligns with clinical expectations: patients with more severe or complex conditions tend to stay longer in the ED for extended evaluation, stabilization, or coordination of downstream care. In contrast, shorter visits are more strongly associated with discharge or leaving without being seen, suggesting that lower acuity or patient-initiated departures occur earlier in the ED encounter. The LOV plot thus reflects the role of clinical acuity and care complexity in shaping final disposition decisions.

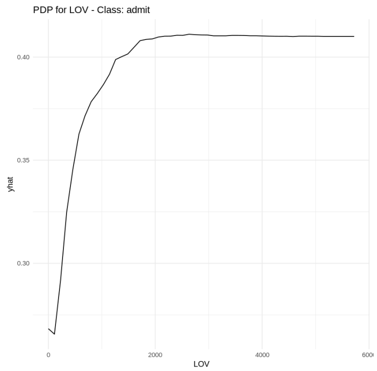
Fig 5 also demonstrates a clinically coherent trend. As pulse increases—an indicator of physiological stress—the model predicts higher probabilities of admission and transfer, consistent with the interpretation that elevated pulse is a marker of critical illness or instability. Conversely, elevated pulse rates significantly reduce the likelihood of discharge, suggesting that patients presenting with tachycardia are less likely to be safely released. Interestingly, the effects on “left” and “returned” are less monotonic, with mild fluctuations across pulse values. This may indicate that these outcomes are influenced by a wider array of non-clinical or contextual factors (e.g., wait times, patient preferences, quality of initial assessment), which cannot be fully explained by vital signs alone.

Together, Fig 4 and Fig 5 show how machine learning models like Random Forest are able to capture complex, non-linear relationships between clinical variables and patient disposition. They also reaffirm that vital signs and care trajectories are core drivers of ED outcomes, and support the interpretability of the model by aligning well with real-world clinical logic.

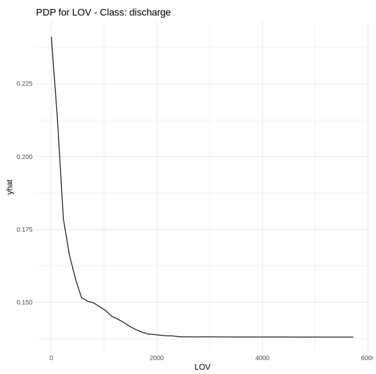
3. Logistic Regression

Class specific Feature Importance Analysis Using LR coefficients: Multinomial logistic regression model with Lasso regularization (LR) gave us a training accuracy of 0.7987, test accuracy of 0.7242 and test AUC score of 0.8726. Despite its modest training performance, LR in fact exhibited the smallest gap between training and test accuracy compared to other models, indicating a decent generalizability. The sparsity that regularization introduced to the model, allowing for a deeper exploration of how the learned coefficients can tell us about the relationships between the patient as well as hospital features and the probabilities of different dispositions.

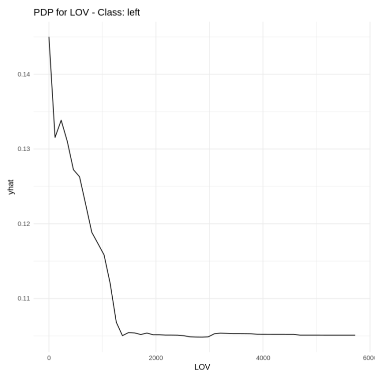
A regularization strength of 0.0027 was chosen through cross-validation, the cv.glmnet () function for multinomial regression has



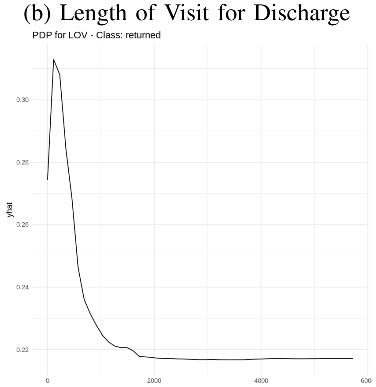
(a) Length of Visit for Admit Patients



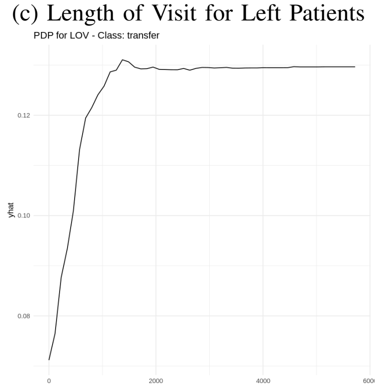
(b) Length of Visit for Discharge



(c) Length of Visit for Left Patients

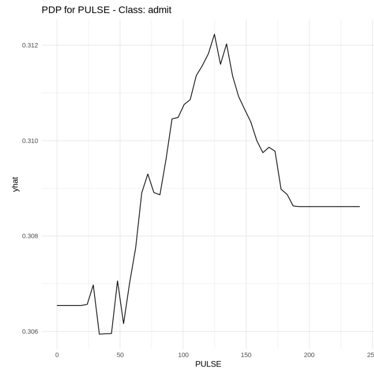


(d) Length of Visit for Return Patients

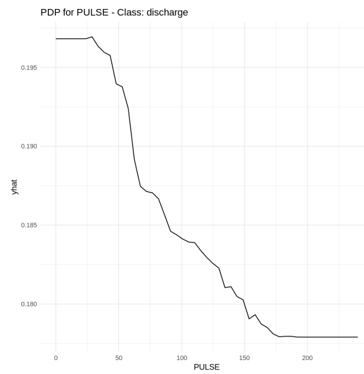


(e) Length of Visit for Transfer Patients

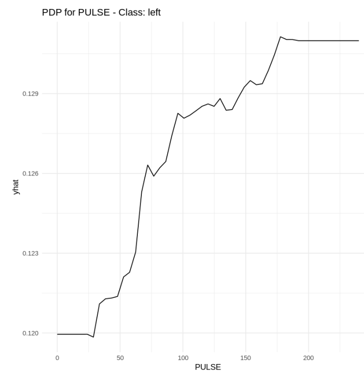
Fig. 4: Overview of Length of Visit by Category



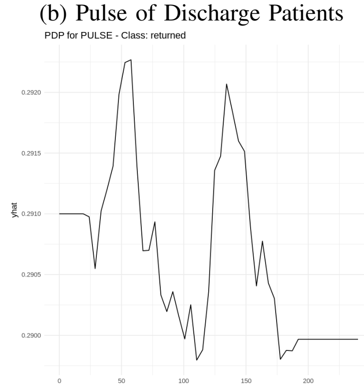
(a) Pulse of Admit Patients



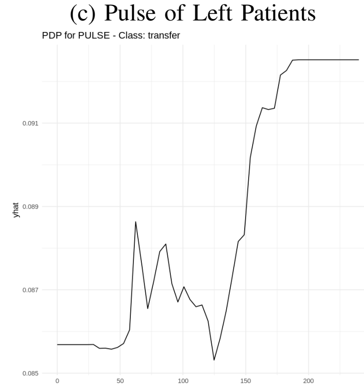
(b) Pulse of Discharge Patients



(c) Pulse of Left Patients



(d) Pulse of Return Patients



(e) Pulse of Transfer Patients

Fig. 5: Overview of Patient's Pulse by Category

a default attribute ‘type.multinomial = “ungrouped”‘, meaning each class has its own set of features and coefficients. We can examine the feature contributions by class when looking into the different subset and number of features with nonzero coefficients.

Across the five disposition classes, each retained approximately 200 features. We grouped these features into the following categories: *Patient Initial Vitals*, *Triage Evaluation*, *Patient Chief Complaint*, *Patient Chronic Disease*, *Patient Injury Report*, *Medical Provider Diagnosis*, *Patient Socioeconomic Characteristics*, *Hospital Usage and Operation*, and *ED Medication and Procedure*. According to Fig 6, The Logistic Regression (LR) model primarily relied on variables from four categories: *Injury*, *Diagnosis*, *Chief Complaint*, and *ED Characteristics*. Together, these accounted for roughly 80% of all retained features. To gain more class-specific insights, we identified the top 20 variables for each disposition class and compared their category distributions.

We found that, across all five classes, variables related to *Patient Injury* were the most influential, comprising about 50% of the top 20 predictors. For admission outcomes, variables from the *Medical Provider Diagnosis* category alone contributed over 50% of the top-ranked features. *Chief Complaint* features were notably important in predicting *Transfer* and *Returned* cases, much like *Procedure*-related

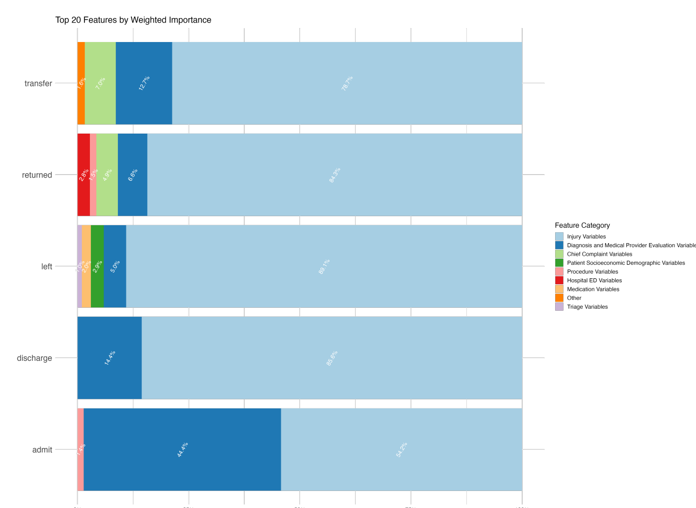


Fig. 6: Top 20 Features by Weighted Importance

variables were informative for both *Admission* and *Returned* outcomes. Additionally, features reflecting *Hospital Usage* patterns appeared relevant for *Returned* cases, while *Medications*, *Demographics*, and *Triage Evaluation* features provided more targeted predictive value for patients who *Left Early*.

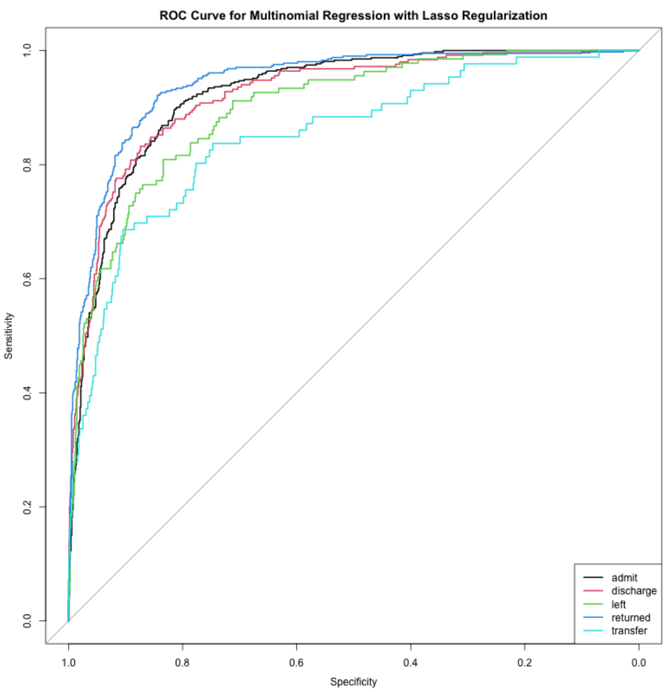


Fig. 7: ROC Curve for Multinomial Regression with Lasso Regularization

According to Fig 7, the ‘returned‘ class was the best classified from our model regarding the test confusion matrix. It would be relevant for a deeper dive into the coefficients of specific variables, such as decoding the injury causes or the diagnosis with ICD code, and look into the chief complaints [7].

Our initial exploration helped us with insights that there are plenty of variables associated with injury caused by accidents such as fire, bitten by wild animals, and have positive and significant coefficients for the ‘returned‘ classification. On the other hand, weaker but still notable predictors such as misdiagnosed COVID-19 and driver injuries from car accidents played more distributed roles across all five classes.

The class-specific and moderately sized feature sets learned by LR helped us with an interpretable mapping from patient and hospital characteristics to ED dispositions. We anticipate further examination of the coefficients will provide insightful understanding of the risk factors behind dispositions.

VII. Discussion

One of the main challenges that we encountered was the imbalance in class support, we addressed this using stratified cross validation for our tree models. Another way to deal with this unbalanced class we would want to try is to use case-control downsampling, by keeping all the cases, or the special dispositions and downsample other control classes.

Another major challenge was the presence of low-variance features in our dataset. While the NHAMCS data offers a rich feature space, many columns are categorical variables that can take on over 500 distinct values. The way we dealt with this issue here was aggregating these distinct values one level up into broader categories, later applying one-hot encoding. However, we ended up with a sparse feature matrix, that is many columns were almost all zero, with one or two one entries. One the one hand this reflects some of the real-world rarity of certain chief complaints, diagnosis, or injury causes, such sparsity poses challenges in our trading process. Looking forward, we are aware of some doable approaches in constructing a better feature space. Such as Chi-squared test for feature selection. Moreover, we also think of treating features like chief complaints as text data and applying text data feature engineering , where we can make the most of the fruitful information that was collected, to anticipate our learning algorithms to make better predictions.

VIII. Conclusion

In this study, we explored how machine learning can help predict what happens to patients after their visit to the emergency room—whether they are admitted to the hospital, discharged, transferred elsewhere, leave before being seen, or return soon after. Using a large national dataset with thousands of emergency visits, we trained several prediction models to identify patterns in patient information and hospital usage.

Among the models tested, **XGBoost** achieved the highest overall accuracy, while **Random Forest** also performed very well—achieving an accuracy of 71.46% and demonstrating strong capability in distinguishing between different patient outcomes. **Logistic Regression**, although simpler, showed the most stable performance between training and testing and was particularly helpful in explaining how specific features influenced decisions. With this model, we found that *injury reports*, *vital signs*, and *diagnosis codes* were especially important in predicting whether a patient would be admitted, transferred, or return soon after discharge.

One of the most important takeaways from our study is that what happens to a patient in the ER is mostly influenced by their current clinical condition—not by who they are. Demographic factors such as race or insurance status had far less predictive power than clinical variables like length of stay, administered medications, and initial vital signs. This finding is encouraging, as it suggests that prediction tools based primarily on medical data can support fairer, faster, and more consistent decisions, ultimately improving patient care and hospital efficiency. In the future, such tools could help hospitals manage bed availability, reduce wait times, and identify high-risk patients early on—all without relying on sensitive or potentially biased personal information.

References

[1] R. Sánchez-Salmerón, J. L. Gómez-Urquiza, L. Albendín-García, M. Correa-Rodríguez, M. B. Martos-Cabrera, A. Velando-Soriano, and N. Suleiman-Martos, “Machine learning methods applied to triage in emergency services: A systematic review,” *International Emergency Nursing*, vol. 60, p. 101109, Jan. 2022.

[2] M. A. Shbool, O. S. Arabeyyat, A. Al-Bazi, A. Al-Hyari, A. Salem, T. Abu-Hmaid, and M. Ali, “Machine Learning Approaches to Predict Patient’s Length of Stay in Emergency Department,” *Applied Computational Intelligence and Soft Computing*, vol. 2023, no. 1, p. 8063846, 2023.

[3] W. S. Hong, A. D. Haimovich, and R. A. Taylor, “Predicting hospital admission at emergency department triage using machine learning,” *PloS One*, vol. 13, no. 7, p. e0201016, 2018.

[4] S. Sezik, M. Ö. Cingiz, and E. İbiş, “Machine Learning-Based Model for Emergency Department Disposition at a Public Hospital,” *Applied Sciences*, vol. 15, no. 3, p. 1628, Jan. 2025.

[5] “Improving the Emergency Department Discharge Process,” https://www.ahrq.gov/patient-safety/settings/emergency-dept/discharge-process.html.

[6] "Ftp.cdc.gov - /pub/Health_Statistics/NCHS/Dataset_Documentation/NHAMCS/stata/," https://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NHAMCS/stata/.

[7] "Ftp.cdc.gov - /pub/Health_Statistics/NCHS/Publications/ICD10CM/2025-Update/," https://ftp.cdc.gov/pub/Health_Statistics/NCHS/Publications/ICD10CM/2025-Update/.

[8] "Estimates of Emergency Department Visits in the United States, 2016-2022," https://www.cdc.gov/nchs/dhcs/ed-visits/index.htm, Jul. 2024.

[9] S. Sezik, M. Ö. Cingiz, and E. İbiş, "Machine Learning-Based Model for Emergency Department Disposition at a Public Hospital," *Applied Sciences*, vol. 15, no. 3, p. 1628, Jan. 2025.

[10] A. Vântu, A. Vasilescu, and A. Băicoianu, "Medical emergency department triage data processing using a machine-learning solution," *Heliyon*, vol. 9, no. 8, p. e18402, Jul. 2023.

[11] "Weekly Provisional Counts of Deaths by State and Select Causes, 2020-2023 | Data | Centers for Disease Control and Prevention," https://data.cdc.gov/NCHS/Weekly-Provisional-Counts-of-Deaths-by-State-and-S/muzy-jte6/about_data.