NHS
Wrightington, Wigan and Leigh
NHS Foundation Trust

Predicting ICU Transfers: A Machine Learning Approach to Early Identification of Critical Patients

Lancaster University

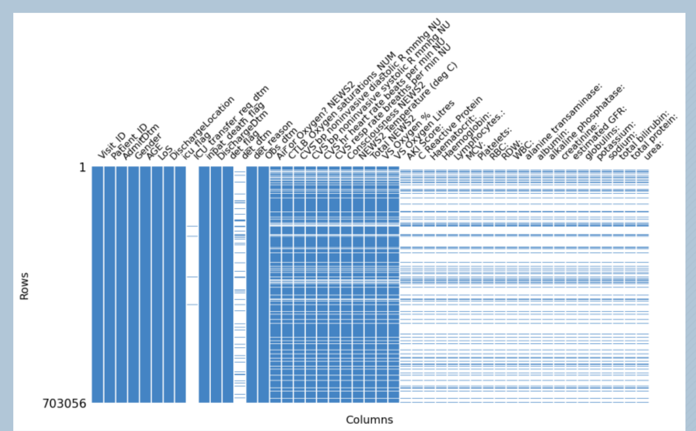
Akash Marar

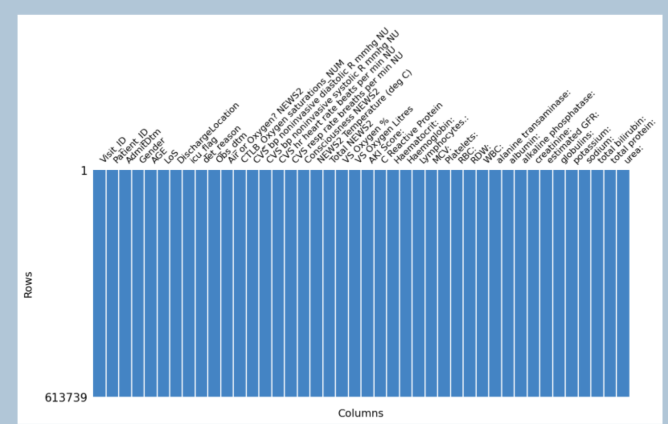
Key Objectives

- Develop a predictive model to accurately forecast ICU transfers using machine learning techniques.
- Address the challenge of class imbalance in ICU transfer data to improve model performance.
- Identify the most important features that contribute to ICU transfer predictions for better clinical decision-making.

Handling Missing Data and Data Correction

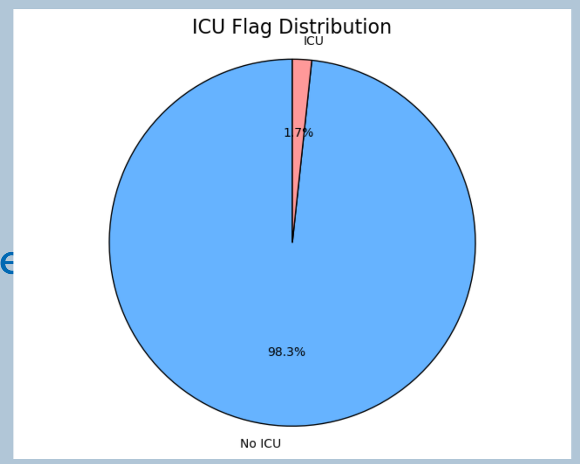
- Data Filtering: Removed rows missing critical timestamp data .
- Imputation: Applied forward and backward filling for time-series data, and iterative imputation to estimate remaining missing values.
- Duplicate and Threshold Filtering: Removed duplicate rows and filtered rows with more than 16 missing values.
- Data Validation: Validated binary variables, corrected outliers, capped percentages, and ensured all values were clinically possible.





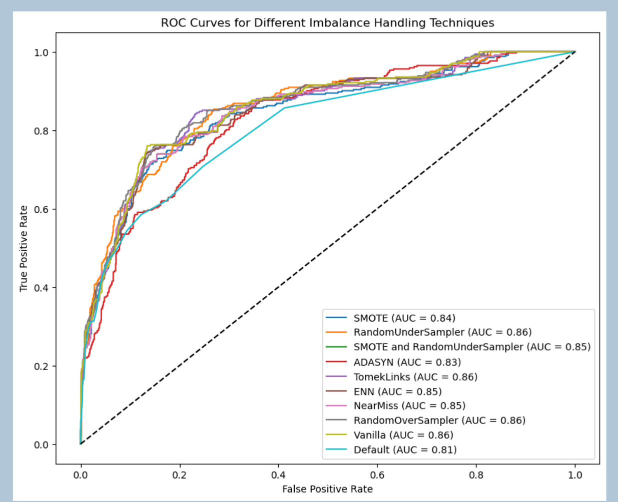
Addressing Class Imbalance

- Only 1.7% of patients required ICU transfer, highlighting a major class imbalance.
- Imbalance makes models biased toward predicting "No ICU" (98.3% majority).
- To improve predictions, various balancing techniques were tested.
- Effective prediction ensures better resource management and improved patient care outcomes.

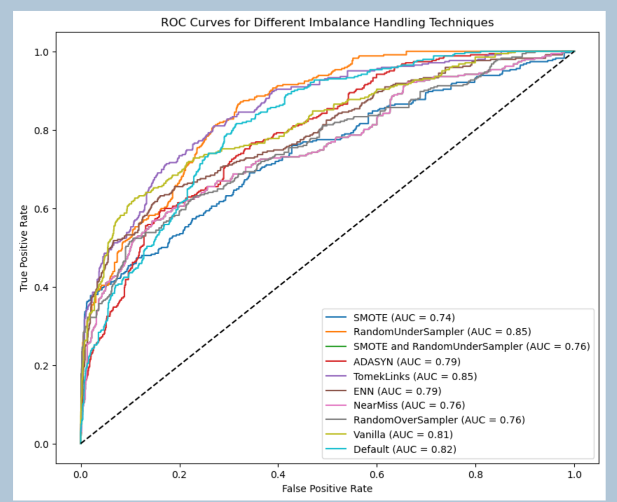


ROC Curve Analysis: Random Forest vs. XGBoost

- **Random Forest** shows more **consistent performance** across different imbalance handling techniques, with AUC values clustering between **0.83 and 0.86**.
- **XGBoost** results are more varied, with AUC values ranging from **0.74 to 0.85**.
- Techniques like **RandomUnderSampler** and **TomekLinks** perform well with both models, but **Random Forest** offers a **more stable ROC** across techniques.



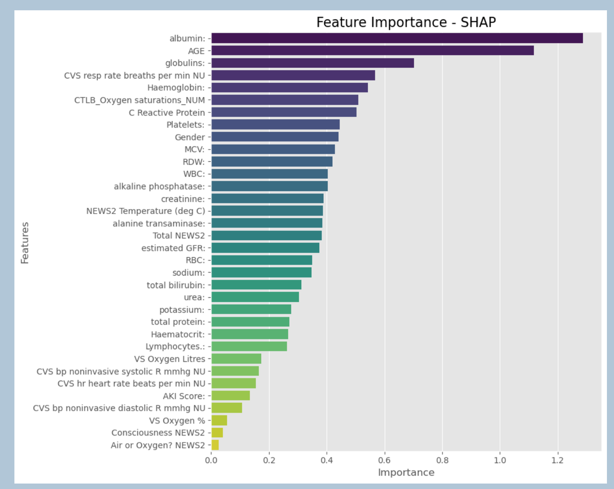
Random Forest



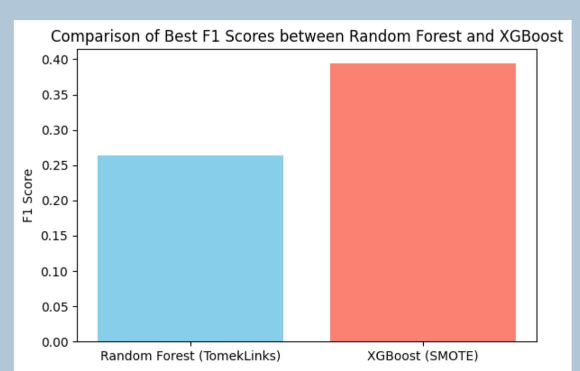
XGBoost

Feature Importance

- Key Features: albumin, age and globulin are the top predictors.
- SHAP Analysis: SHAP reveals the non-linear impact of features like age, which ranks higher than in traditional methods.
- Balanced Data: Vital signs and lab values both play crucial roles in ICU transfer predictions.
- Clinical Relevance: Combining easily measurable bedside data with lab results provides a balanced and actionable model for ICU transfer predictions.



F1-Score and Best Performing Models



- While ROC curves show that Random Forest consistently performed better, we chose the F1-score as our primary metric.
- The F1-score was selected to balance precision (avoiding unnecessary ICU transfers) and recall (capturing patients in need of ICU care)
- The best results were achieved using XGBoost with SMOTE and Random Forest with TomekLinks.
- XGBoost with SMOTE delivered the highest F1 score, making it the most effective model

Final Remarks

- XGBoost with SMOTE outperformed the traditional NEWS2 score in ICU tra nsfer prediction.
- SHAP provided valuable insights into key features, enhancing the interpretability of the model.
- Future work will focus on improving model performance through additional data such as diagnostic notes
- Also exploring the use of other machine learning algorithms and more complex balancing techniques to improve the output

akashmarar@gmail.com