

ICU Tableau

Title	Introduction	Research Question	Main Objectives	Objective 1	Objective 2	Objective 3

Predictive Models In ICU Databases

Alain, Darlene, Mariah

DATA 4381: DATA CAPSTONE

The University of Texas at Arlington
College of Science - Division of Data Science

ICU Tableau

Title	Introduction	Research Question	Main Objectives	Objective 1	Objective 2	Objective 3

What is the Intensive Care Unit (ICU)?

The ICU is a specialized hospital unit that provides continuous monitoring and life-support for patients with critical or life-threatening conditions. Individuals admitted to the ICU often experience rapidly changing conditions that require immediate medical attention and constant observation.

What problems does ICU data have?

ICU data presents several challenges for clinical analysis and machine learning. One major issue is class imbalance because severe outcomes such as death or hospital readmission occur far less frequently than positive outcomes. Because most patients survive or recover, the data is dominated by the "good outcome" class. This imbalance causes predictive models to overlook or poorly identify the minority class (the patients who actually need early risk detection the most).

Why does this matter?

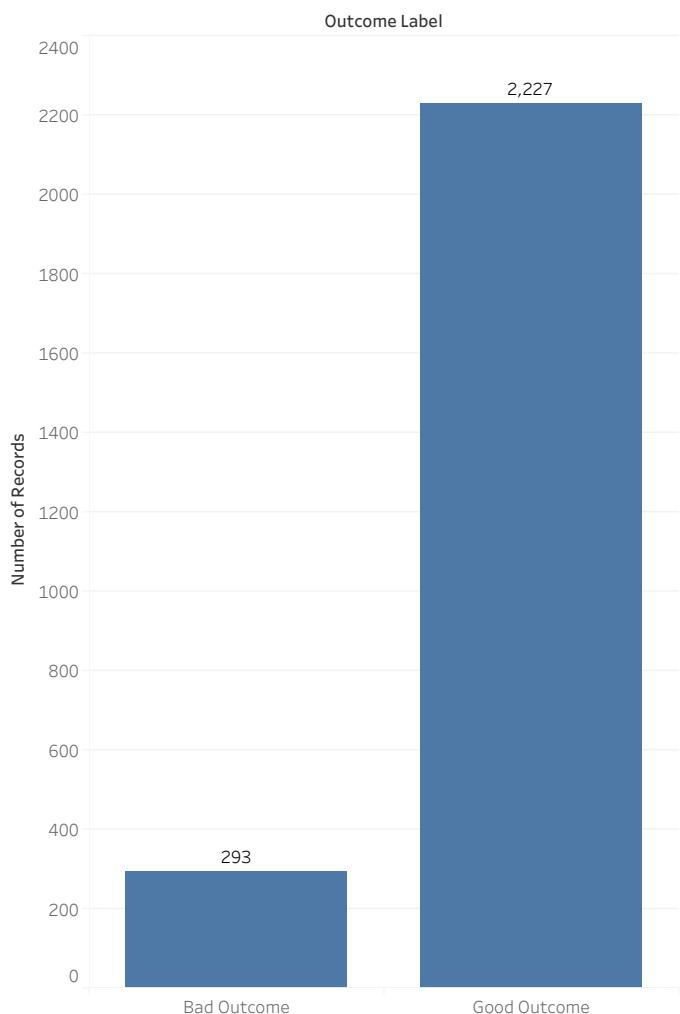
Identifying high-risk ICU patients early can improve care decisions, resource allocation, and patient survival.

Overview of Dataset

The eICU database provides real ICU patient data with vitals, labs, demographics, and outcomes.

Link: <https://eicu-crd.mit.edu/>

Class Imbalance



ICU Tableau

Title	Introduction	Research Question	Main Objectives	Objective 1	Objective 2	Objective 3

Research Question:

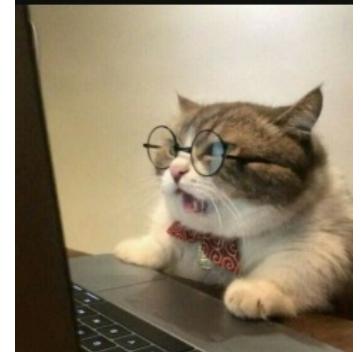
Can we **predict ICU patient poor outcomes (death or readmission)** using demographic, clinical, and physiological data from the dataset, and how does imbalance affect model performance?

ICU Tableau

Title	Introduction	Research Question	Main Objectives	Objective 1	Objective 2	Objective 3

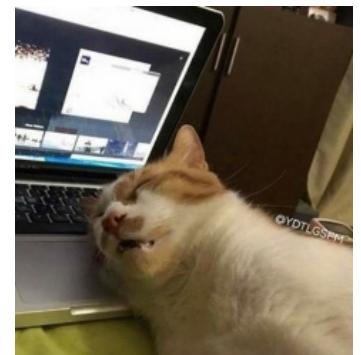
Objective 1 (Mariah)

Is training on an imbalanced dataset really better than training on a balanced dataset?????



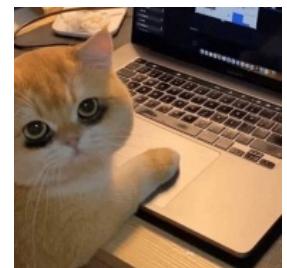
Objective 2 (Alain)

What things are sacrificed if we wanted to make a model that was completely based on real-life prevalence?????



Objective 3 (Darlene)

What things are sacrificed if we wanted to make a model that was completely optimized to its best performance? ?? ?



ICU Tableau

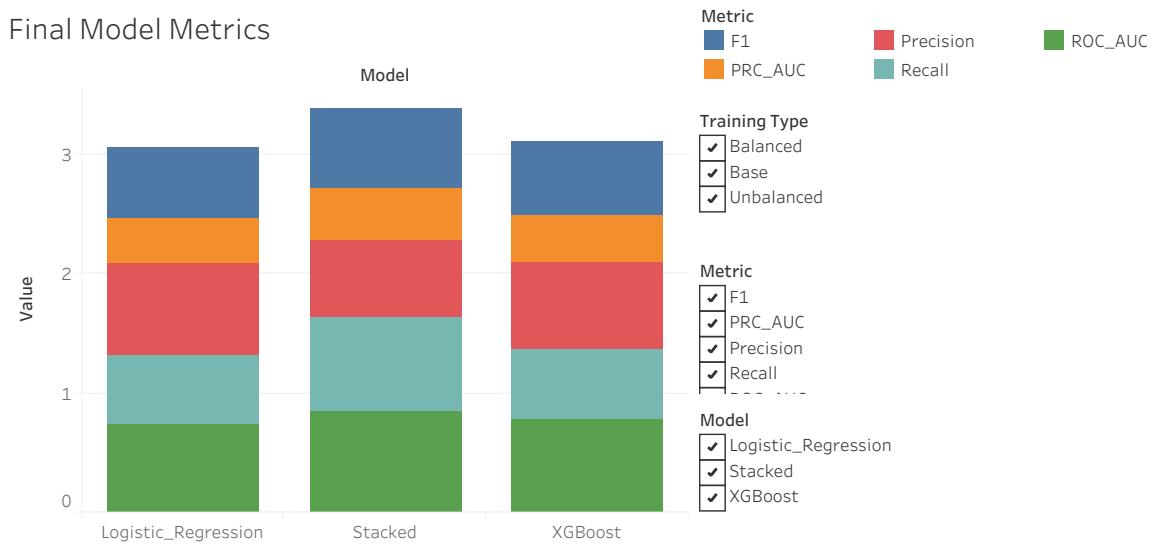
Introduction	Research Question	Main Objectives	Objective 1	Objective 2	Objective 3	Conclusion
--------------	-------------------	-----------------	-------------	-------------	-------------	------------



OBJECTIVE 01

Investigate the impact of class imbalance by testing multiple data scenarios, including class weighting and the original outcome distribution

Final Model Metrics



Research Question	Main Objectives	Objective 1	Objective 2	Objective 3	Conclusion	Thank You
-------------------	-----------------	-------------	-------------	-------------	------------	-----------

OBJECTIVE 02

Construct a clinically realistic model that reflects real-world constraints, prioritizing interpretability and practicality over raw metrics.



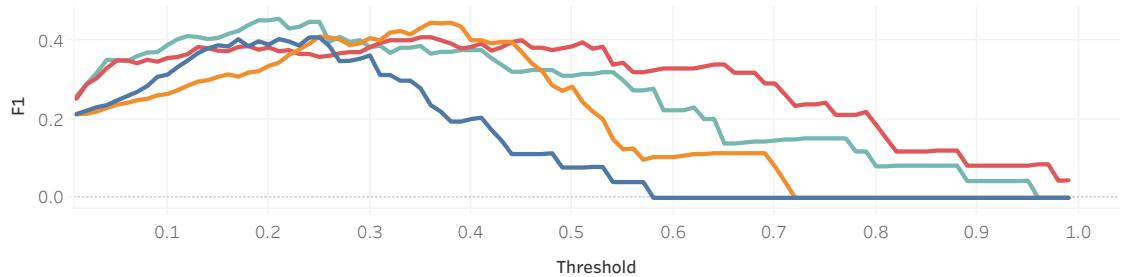
ICU Tableau

Research Question	Main Objectives	Objective 1	Objective 2	Objective 3	Conclusion	Thank You
-------------------	-----------------	-------------	-------------	-------------	------------	-----------



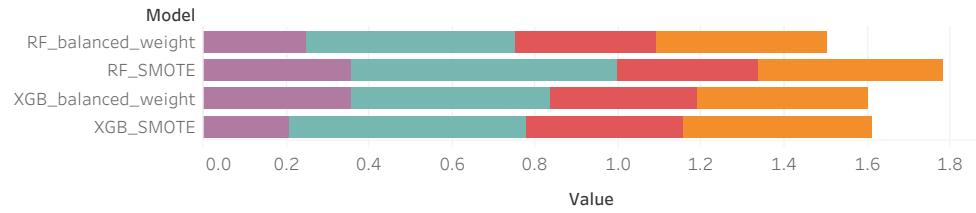
Model
RF_balanced_weight
RF_SMOTE
XGB_balanced_weight
XGB_SMOTE

F1 vs Threshold



Measure Names
F1
Precision
Recall
threshold

SMOTE & Class Weight Model Metrics



ICU Tableau

Research Question	Main Objectives	Objective 1	Objective 2	Objective 3	Conclusion	Thank You

Summarized Findings

Mariah:

- ~ Training on an imbalanced or balanced dataset depends on your research question
- ~ The balanced model will usually perform better, but that comes with trade-offs

Alain:

Darlene:

- ~ ICU mortality prediction **cannot rely on default thresholds** because rare outcomes shift optimal decision points
- ~ All models achieved **higher F1 scores at lower thresholds**, showing that threshold tuning is necessary
- ~ **SMOTE** improves minority-class detection and significantly boosts **recall**

Future Work

- ~ Create real-life scenario model
- ~ Create dashboard for this data
- ~ Improve stacked model
- ~ Use common techniques used on medical data, such as Multiple Iteration

Research Question	Main Objectives	Objective 1	Objective 2	Objective 3	Conclusion	Thank You

Contact Us!



Alain Areeba Siddiqui
axs6903@mavs.uta.edu



Mariah Noelle Cornelio
mnc3287@mavs.uta.edu



Darlene Eligado
dae9134@mavs.uta.edu