Problem: Sometimes the hospital capacity is full and it needs to alleviate pressure by either more staff, or more resources. However it can also cut the demand if it can determine that an encounter does not need further care.

Solution: To help hospital address the problem we can help them find some solution by answering the question below:

1) If a hospital need to know how long until the next available critical care unit is available what data do they need to know this?
   - They need to know when the next ICU or OR is available
2) If a room is occupied, what will help us know ahead when it will be available
   - We can use details about the encounter with the patient and historical data to estimate how long an encounter might take for critical care

Prepping Data - Preprocessing

Step #1: Replacing missing data with nan

Step #2: Remove any columns having more than 4000 nan

Step #3: remove all remaining row that have nan in any column

Step #4: Encode categorical data (readmitted ['NO', '<30', '>30'] => [0, 1]), labelEncode (gender, race)

Step #5: Doing a correlation matrix to remove any features that have high correlation together (aka Feature Selection)

Step #6: Splitting into training data and testing data

Model Selection | Training | First sight Evaluation

Selection

- KNN, Decision Tree, Logistic Regression (probability / classification), Random Forest(classification)

First sight evaluation

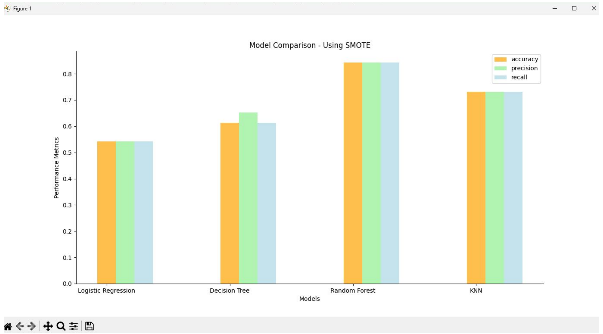
- All had high negative prediction (no readmission) accuracy rate
- All had supper low positive prediction accuracy rate
- We try to hyperparameter tuning our model using shuffle and cross validation a technique to check for different model properties to find best performing model variate
- Reason is that the data set had 90000 points, but 80000 were negative cases, so the model was bias toward negative cases. So we use SMOTE a technique for sampling to create synthetic positive cases

```
Model: RandomForestClassifier
              precision    recall  f1-score   support

           0       0.89      0.97      0.93     21790
           1       0.13      0.04      0.06      2724

    accuracy                           0.86     24514
   macro avg       0.51      0.50      0.49     24514
weighted avg       0.81      0.86      0.83     24514

C:\Users\Danh1\AppData\Local\Programs\Python\Python311\Li
fitted without feature names
  warnings.warn(
Model: KNeighborsClassifier
              precision    recall  f1-score   support

           0       0.89      0.99      0.93     21790
           1       0.12      0.01      0.03      2724

    accuracy                           0.88     24514
   macro avg       0.50      0.50      0.48     24514
weighted avg       0.80      0.88      0.83     24514

C:\Users\Danh1\AppData\Local\Programs\Python\Python311\Li
tted without feature names
  warnings.warn(
Model: LogisticRegression
              precision    recall  f1-score   support

           0       0.92      0.05      0.09     21790
           1       0.11      0.97      0.20      2724

    accuracy                           0.15     24514
   macro avg       0.52      0.51      0.15     24514
weighted avg       0.83      0.15      0.10     24514
```

First Sight Evaluation Continue…

Next Steps:

- Evaluate model performance of KNN, Logistic Regression, Decision Tree, Random Forest
- Feature Importance
- Hyperparameter Tuning using GridsearchCV (cross validation)



Using Neural Network

- Initially was poor because we misplaced the column name causing the model to wildly misinterpret the data. This was because we wanted to predict readmitted, but readmitted column label was given to the wrong column in the data causing low accuracy at 60%

Application of these models

I. Next available unit

Look at the inventory of resources (ICU | OR), if they are all occupied, then run the model on each encounter that are occupying the space, and find the encounter with the earliest discharged disposition.

Output: discharged disposition in days

II. Ready to be discharged?

Looking at the encounter of the patients and the decision of the doctor to discharge them, the model can check to confirm if the decision to discharge is one that will not produce a within 30-days readmission

Output: 30-days readmission?