The timing of patient liberation from ventilatory support is critical to avoid complications and reduce healthcare costs. Models that accurately predict independence from mechanical ventilation would be pivotal in supporting clinician’s informed decision-making, improving patient outcomes upon extubation, and decreasing ICU expenditures.

### Aim 1: Data Preprocessing
- Inclusion/exclusion criteria
- Data Imputation and cleaning
- Sliding window definition

### Aim 2: Liberation Prediction
- Predict whether the patient will have an extubation success or failure based on the status

### Aim 3: Duration Prediction
- Predict whether the patient can be extubated within a week based on the sliding window
- Predict whether the patient can be extubated within which quantile range

### DATASET
- 9781 patients with multiple 6-hour observation windows were pulled from the Precision Medicine Analytics Platform (PMAP) Dataset feature summary is shown below.
  - Heterogenous data of multiple modalities: demographic, vital signs, lab results, medical history
  - Predominantly consists of overweight, geriatric, white males with cardiovascular risk factors and cystic fibrosis

### Table 1: Feature summary table

<table>
<thead>
<tr>
<th>Feature</th>
<th>Type</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>int</td>
<td>18-100</td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td>float</td>
<td>18-45</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>bool</td>
<td>0: male, 1: female</td>
<td></td>
</tr>
<tr>
<td>Race</td>
<td>float</td>
<td>0: black, 1: other</td>
<td></td>
</tr>
<tr>
<td>Hospital</td>
<td>str</td>
<td>1-100</td>
<td></td>
</tr>
<tr>
<td>Vitals</td>
<td>float</td>
<td>18-100</td>
<td></td>
</tr>
<tr>
<td>Lab</td>
<td>float</td>
<td>18-100</td>
<td></td>
</tr>
</tbody>
</table>

### METHOD
- **Data Imputation & Cleaning Details**
  - Data Imputation: Used KNN imputation with k = 10
  - Selecting successfully extubated patients: Patients successfully extubated if they were not reintubated with 72 hours of the extubation event
  - Data Cleaning: Removed erroneous values, extreme outliers, etc. from the data

- **Final Dataset 21985 windows**
  - Train: Validation: Test = 17709:1978:2298 windows

- **Apply Machine learning & Deep learning Models**
  - Machine Learning Models:
    - Logistic Regression (LR)
    - Random Forest (RF)
    - XGBoost
    - Deep Learning Models:
      - Transformer
      - Recurrent Neural Network (RNN)
      - Long Short-Term Memory Network (LSTM)
      - Convolutional Neural Network (CNN)

- **Hyperparameter Tuning**
  - Based on the train and validation data set

### CONCLUSION
- We created the data set from PMAP for the prediction of patient’s extubation time using 6-hour observation windows, which consists of 9781 patients with multiple features.
- We compared 3 machine learning models and 4 deep learning models on the data set and found XGBoost performed the best on both binary and quantile prediction.
- We conducted SHAP analysis using the trained XGBoost model. We found that lab data is highly relevant to the prediction of extubation time in addition to vital data like blood pressure and Glasgow coma scale. We found platelets in blood, ratio of nucleated erythrocytes to 100 leukocytes, and erythrocyte distribution width are the top 3 most important features in the prediction of extubation time.
- Future work can be done in creating a classifier that outputs a continuous time-forecast for the end of a mechanical ventilation instance.

### REFERENCES