



Track patient recovery in real-time by processing streaming data

BIOMEDICAL DATA DESIGN

TA: Haoyin Xu

Group: Zhenyu Xiao
Haobin Zhou
Yimeng Xu
Emma Cardenas

The slide features a white background with a black border. In the corners, there are decorative blue circles: a large one in the top-left, a medium one in the top-right, and a small one in the bottom-left.

01

Model testing (with flattened data)

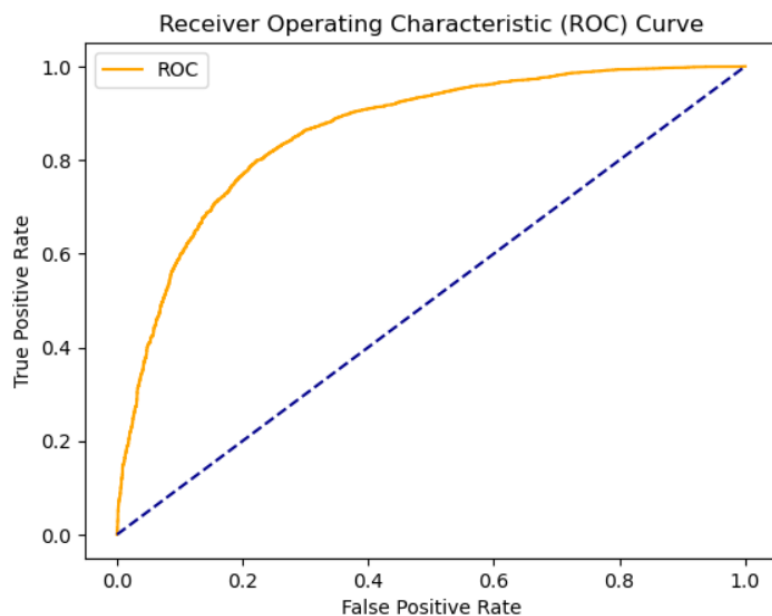
01 Model testing

Logistic regression

```
accuracy: 78.73%
confusion_matrix:
[[2973  865]
 [ 768 3070]]
classification_report:
              precision    recall  f1-score   support

     0       0.79       0.77       0.78       3838
     1       0.78       0.80       0.79       3838

 accuracy          0.79          0.79          0.79       7676
 macro avg         0.79          0.79          0.79       7676
 weighted avg      0.79          0.79          0.79       7676
```

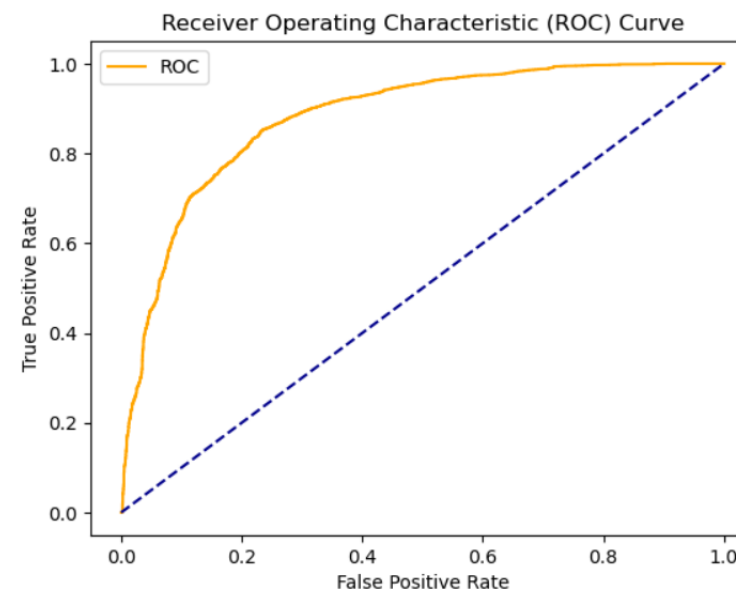


Random forest

```
accuracy: 80.82%
confusion_matrix:
[[3052  786]
 [ 686 3152]]
classification_report:
              precision    recall  f1-score   support

     0       0.82       0.80       0.81       3838
     1       0.80       0.82       0.81       3838

 accuracy          0.81          0.81          0.81       7676
 macro avg         0.81          0.81          0.81       7676
 weighted avg      0.81          0.81          0.81       7676
```



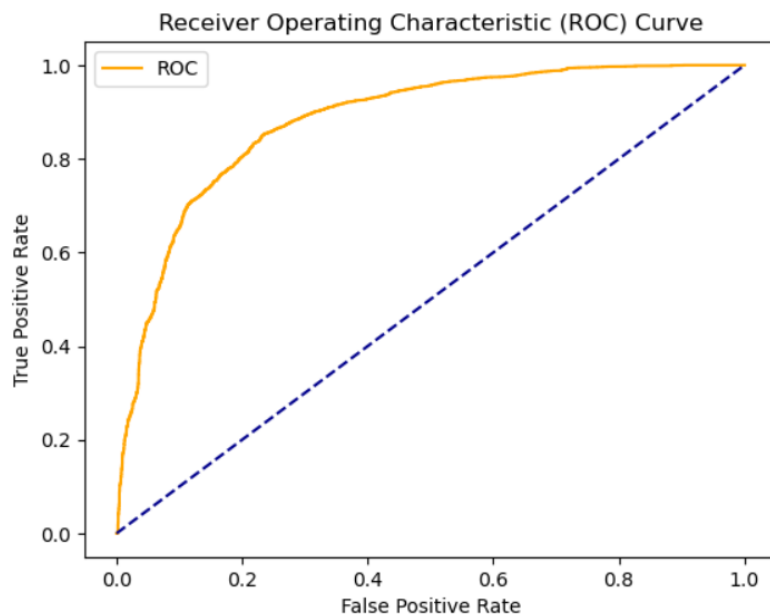
01 Model testing

XGboost

```
accuracy: 80.91%
confusion_matrix:
[[3082  756]
 [ 709 3129]]
classification_report:
              precision    recall  f1-score   support

     0           0.81       0.80       0.81       3838
     1           0.81       0.82       0.81       3838

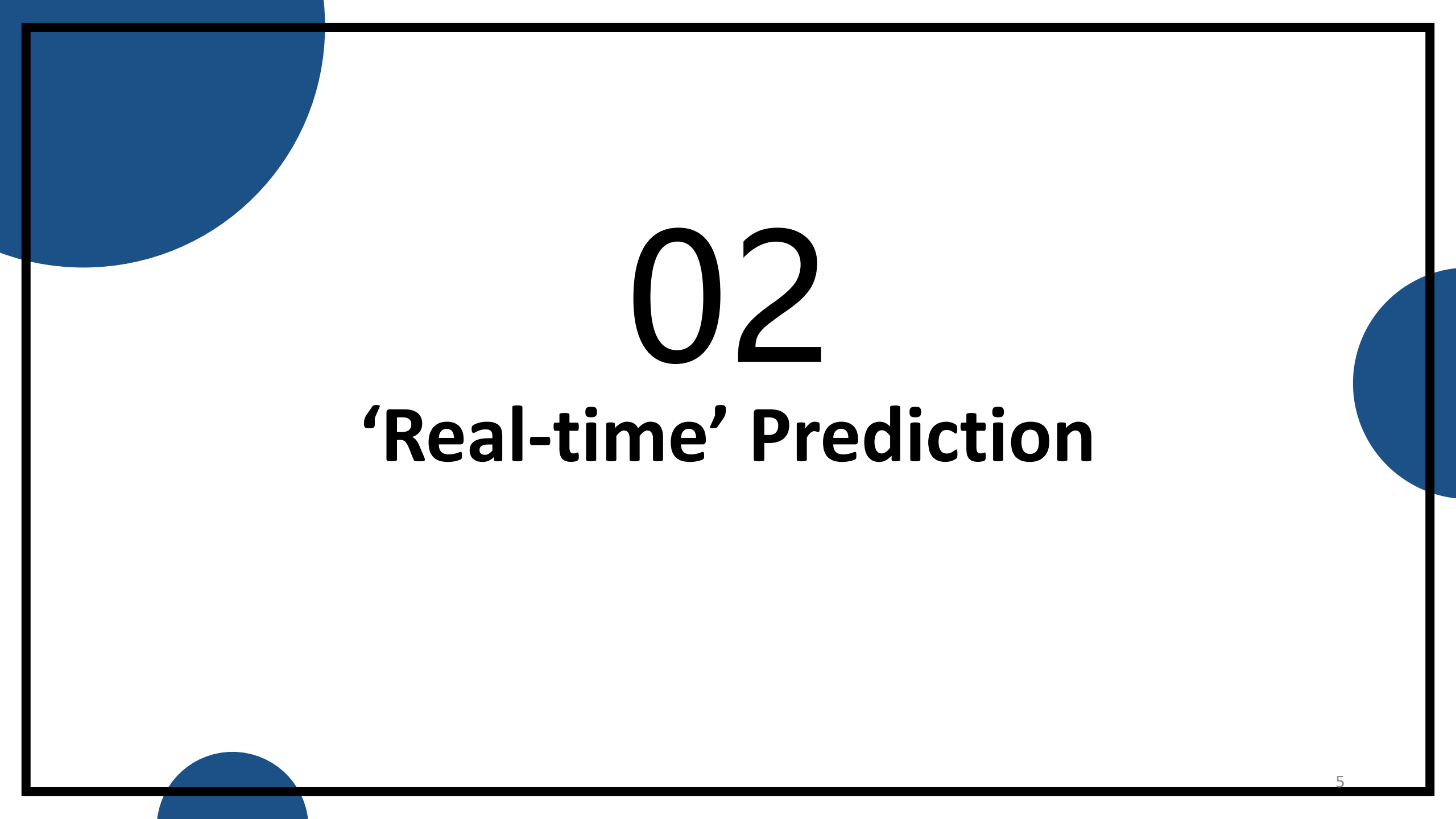
 accuracy          0.81          0.81          0.81       7676
  macro avg          0.81          0.81          0.81       7676
 weighted avg          0.81          0.81          0.81       7676
```



Accuracy:~80%

Next steps:

1. Improve the parameters
2. How to achieve 'real-time' prediction?
3. Other features?

The slide features a white background with a black border. In the corners, there are blue circular shapes: a large one in the top-left, a medium one in the top-right, and a small one in the bottom-left.

02

‘Real-time’ Prediction

02 'Real-time' Prediction

Predict the final mortality

↓ We want

Predict the mortality in one day

↓ We need

Remake labels to indicate whether the patients die in one day
(based on the length of data we use)

1. Send first day data to get the prior of the second day
2. Use the first day results to update the state of model: RNN

02 'Real-time' Prediction

Check the number of patients

```
df = pd.read_csv('filtered_low_features_over2days.csv')
unique_patient_count = df['patientunitstayid'].nunique()
print("Number of patients who have more than 48 hours:", unique_patient_count)
```

✓ 1.6s

Number of patients who have more than 48 hours: 5741

```
unique_patient_count_mortality_0 = merged_df1[merged_df1['actualicumortality'] == 0]['patientunitstayid'].nunique()
unique_patient_count_mortality_1 = merged_df1[merged_df1['actualicumortality'] == 1]['patientunitstayid'].nunique()

print("Patients who die:", unique_patient_count_mortality_0)
print("Patients who are alive:", unique_patient_count_mortality_1)
```

✓ 0.0s

Patients who die: 796

Patients who are alive: 4945

The slide features a white background enclosed by a thick black rectangular border. Three large, solid blue circles are positioned at the corners: one in the top-left, one in the bottom-left, and one on the right side. The text "Thank you" is centered in the middle of the slide.

Thank you