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

01

Model testing

01 Model testing

```
accuracy: 72.01%
confusion_matrix:
[[4586 2180]
 [1674 5331]]
classification_report:
              precision    recall  f1-score   support

           0       0.73      0.68      0.70      6766
           1       0.71      0.76      0.73      7005

    accuracy                           0.72     13771
   macro avg       0.72      0.72      0.72     13771
weighted avg       0.72      0.72      0.72     13771
```

Logistic regression

```
accuracy: 72.79%
confusion_matrix:
[[4627 2139]
 [1608 5397]]
classification_report:
              precision    recall  f1-score   support

           0       0.74      0.68      0.71      6766
           1       0.72      0.77      0.74      7005

    accuracy                           0.73     13771
   macro avg       0.73      0.73      0.73     13771
weighted avg       0.73      0.73      0.73     13771
```

Random forest

01 Model testing

```
accuracy: 69.18%
confusion_matrix:
[[4395 2371]
 [1873 5132]]
classification_report:
              precision    recall  f1-score   support

           0       0.70      0.65      0.67      6766
           1       0.68      0.73      0.71      7005

    accuracy                           0.69     13771
   macro avg       0.69      0.69      0.69     13771
weighted avg       0.69      0.69      0.69     13771
```

XGboost

Flatten features & Fix low frequency features (ex. lab)

High frequency features (HR, BP, Temp) to generate results

Use min, max, variance for the hour samples

02

Reprocess the data

02 # Reprocess the data

	patientunitstayid	observationoffset	heartrate	temperature	paO2	FiO2_x	FiO2_y	Glasgow score
0	141515.0	0	86.823529	36.900429	108.5	69.034181	69.034181	6.689885
1	141515.0	1	90.222222	36.958959	283.0	100.000000	100.000000	6.578818
2	141515.0	2	91.600000	36.584615	108.5	68.208171	68.208171	6.462798
3	141515.0	3	87.875000	38.775000	55.0	50.000000	50.000000	6.342544
4	141515.0	4	85.875000	38.362500	71.0	70.000000	70.000000	6.218848

BP	BUN	WBC x 1000	bicarbonate	sodium	potassium	total bilirubin	actualicumortality
87.272727	75.049159	11.7	20.589048	133.145147	3.600000	1.2	1
111.000000	74.894458	11.7	21.000000	132.000000	3.600000	1.2	1
120.800000	74.731208	11.7	20.553918	133.426232	3.704932	1.2	1
111.062500	74.559406	11.7	20.540583	133.569442	3.876114	1.2	1
115.333333	74.379056	11.7	20.530217	133.714358	4.000000	1.2	1

Almost not change

02 Reprocess the data

Heartrate

	patientunitstayid	observationoffset	heartrate	heartrate_max	heartrate_min	heartrate_var	unitdischargeoffset
0	151900.0	0	88.153846	106.0	68.0	132.974359	57
1	151900.0	1	75.500000	88.0	68.0	36.454545	57
2	151900.0	2	71.333333	86.0	64.0	47.238095	57
3	151900.0	3	84.125000	96.0	74.0	45.583333	57
4	151900.0	4	78.625000	92.0	66.0	62.783333	57

Temperature

	patientunitstayid	observationoffset	temperature	temp_max	temp_min	temp_var	unitdischargeoffset
0	151900.0	-5	36.9	36.9	36.9	NaN	57
1	151900.0	0	36.7	36.7	36.7	NaN	57
2	151900.0	2	36.5	36.5	36.5	NaN	57
3	151900.0	6	36.9	36.9	36.9	NaN	57
4	151900.0	10	37.1	37.1	37.1	NaN	57

Should be zero

Thank you