Track patient recovery in real-time by processing streaming data

BIOMEDICAL DATA DESIGN

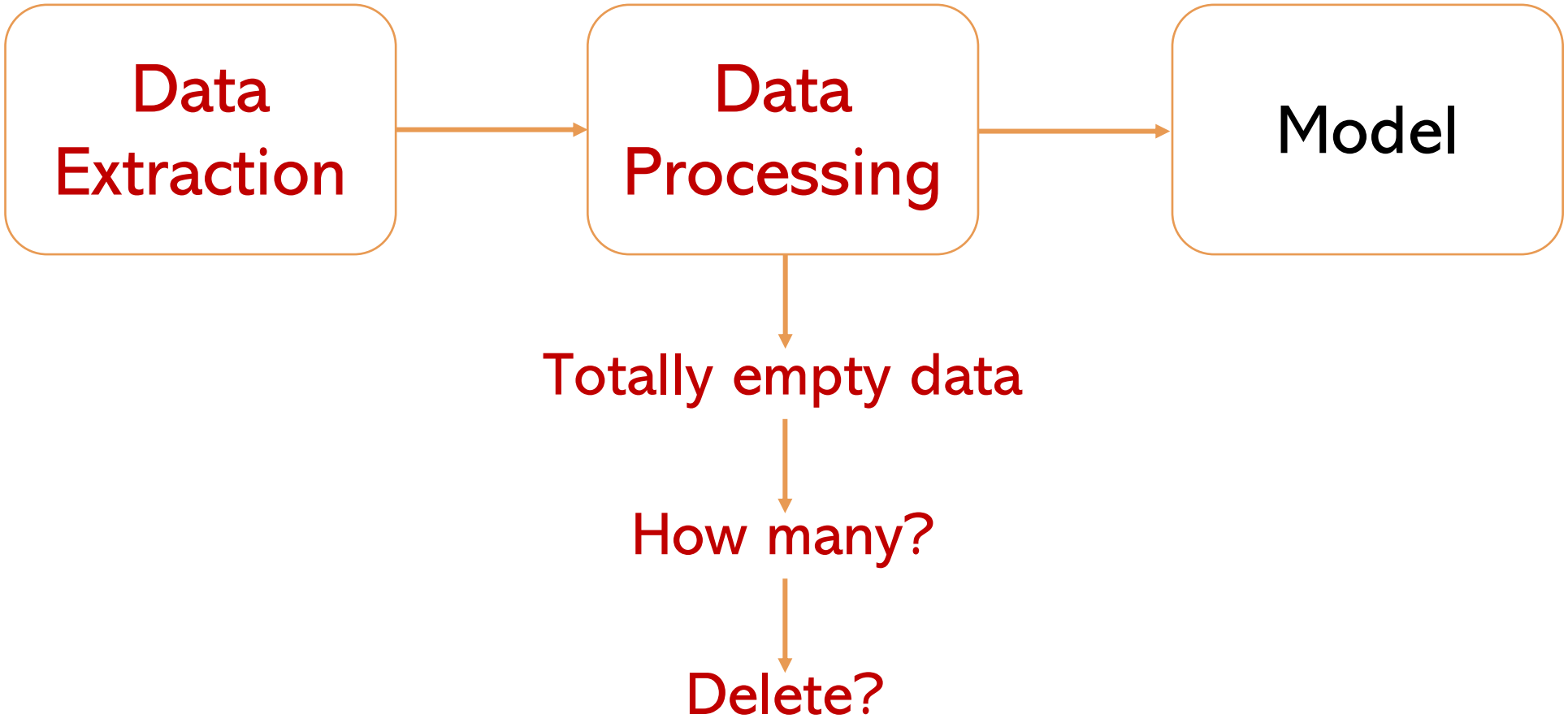
TA: Haoyin Xu

Group:Zhenyu Xiao
Haobin Zhou
Yimeng Xu
Emma Cardenas

01

Data processing

01 Data processing



01

Data processing

feature	unit	item-name	table
Glasgow coma scale		GCS Total	Nursecharting
		Glasgow coma score	Nursecharting
Systolic blood pressure（mmHg）	mmHg	Noninvasivesystolic	Avtalperiod
	mmHg	Non-Invasive BP Systolic	Nursecharting
	mmHg	Invasive BP Systolic	Nursecharting
	mmHg	Systemicsystolic	Vitalperiod
Heart rate		Heart Rate	Nursecharting
Body temperature（≥39ºC (102.2ºF)）	ºC	temperature	Vitalperiod
	ºC	Temperature(C)	Nursecharting
pao2/fio2 ratio（mm Hg/%）			
	mmHg	paO2	lab
	%	FiO2	lab
	%	FiO2	respiratoryCharting
	%	FiO2(%)	respiratoryCharting
Urine output（mL/day）	mL/day	Outputtotal	IntakeOutput
		intakeoutputoffset	IntakeOutput
Serum urea nitrogen level（mg/dL）	mg/dL	BUN	lab
White blood cells count（x 10^3/mm^3）	K/mcL	WBC x 1000	lab
Serum bicarbonate level（mEq/L）	mmol/L	bicarbonate	lab
Sodium level（mEq/L or mmol/L）	mmol/L	sodium	lab
Potassium level（mEq/L）	mmol/L	potassium	lab
Bilirubin level（mg/dL）	mg/dL	total bilirubin	lab

01 Our progress

```
patient_id_all = df_wanted['patientunitstayid'].unique()
patient_id_all.sort()
print(patient_id_all)

# import patient.csv
df_patient = pd.read_csv('patient.csv')
df_patient.sort_values(by=['patientunitstayid'], inplace=True)
df_patient_buf = df_patient[df_patient['patientunitstayid'].isin(patient_id_all)]
df_1time_patient = df_patient_buf[df_patient_buf['unitvisitnumber']==1]
# print(df_1time_patient)

# select the patient id from df_1time_patient
patient_id = df_1time_patient['patientunitstayid'].unique()
print(f'Total number of patients: {len(patient_id)}')
```

```
[ 141168  141203  141227 ... 3353216 3353235 3353251]
Total number of patients: 71353
```

```
# 第一次进入ICU的心血管病疾病患者ID集合
first_time_icu_patients = set(patient_id)
print(f'Number of first time ICU patients: {len(first_time_icu_patients)}')

# 有heartrate数据的患者ID集合
patients_with_heartrate = set(HR['patientunitstayid'].unique())
print(f'Number of patients with heart rate data: {len(patients_with_heartrate)}')
```

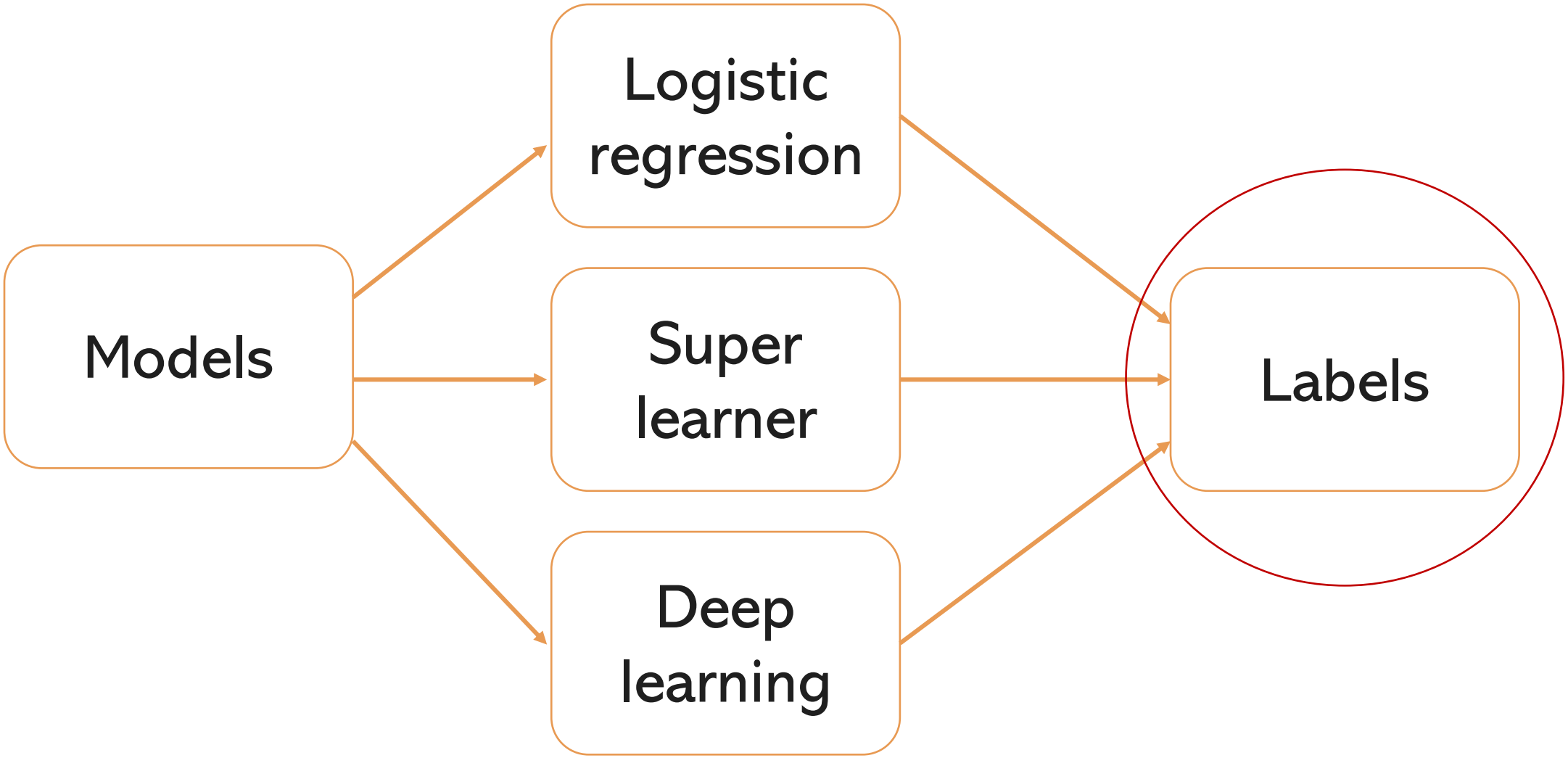
```
Number of first time ICU patients: 71353
Number of patients with heart rate data: 71242
```

```
Final available patients: 15834
```

02

Label

02 Label


Models
Logistic regression
Super learner
Deep learning
Labels

02 Label

```
mortality_df['actualicumortality'] = mortality_df['actualicumortality'].apply(lambda x: 0 if x == 'EXPIRED' else 1)
mortality_df
```

	patientunitstayid	actualicumortality
0	141168	0
2	141194	1
4	141203	1
6	141208	1
8	141227	1
...	...	...
297054	3353235	1
297056	3353237	1
297058	3353251	1
297060	3353254	1

Thank you