



# 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 blue circular shapes: a large one in the top-left, a medium one in the top-right, a small one in the bottom-left, and a medium one in the bottom-right.

# 01

## **Data processing**

## 02 Our progress



## 02 Aligning data

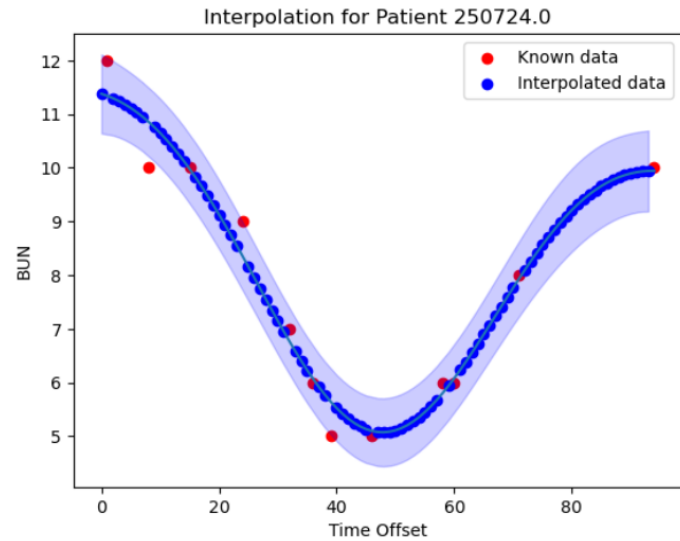
		Glasgow coma score	Nursecharting
Systolic blood pressure (mmHg)	mmHg	Noninvasivesystolic	Avtalperiod
	mmHg	Non-Invasive BP Systolic	Nursecharting
	mmHg	Invasive BP Systolic	Nursecharting
	mmHg	Systemicsystolic	Vitalperiod
		Pasystolic	Vitalperiod
Heart rate		Heart Rate	Nursecharting
	°C	temperature	Vitalperiod
Body temperature (≥39°C (102.2°F))	°C	Temperature(C)	Nursecharting
	°F	Temperature (F)	Nursecharting
		Temperature Location	Nursecharting
	%	FiO2	lab
	mmHg	paO2	lab
pao2/fio2 ratio (mm Hg/%)		SVO2	Nursecharting
		O2 Saturation	Nursecharting
	%	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 <sup>3</sup> /mm <sup>3</sup> )	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

```
def align_data(patient_batch, patient_offset, data, kernel='C(1.0) * RBF(10) + WhiteKernel(noise_level=1, noise_level_bounds=(1e-10, 1e5))'):
    """
    Summary: align data and interpolate missing values

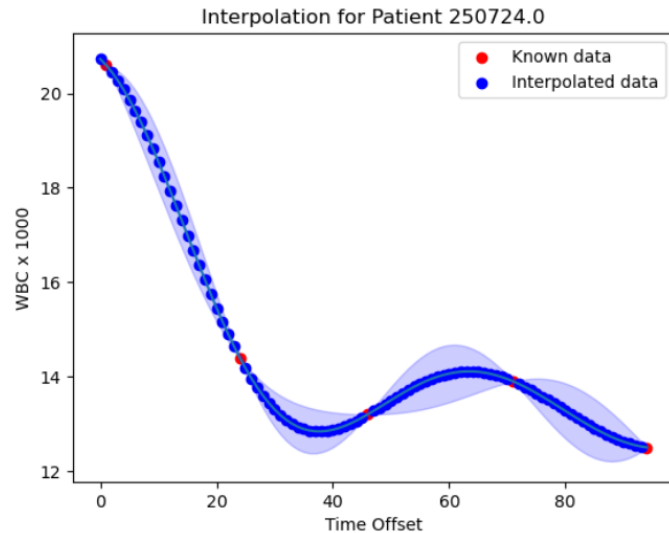
    Args:
        patient_batch: the list of wanted patient id, used to split data
        patient_offset: the dataframe of patient offset data, including patientunitstayid, unitdischargeoffset
        data: the dataframe of data, including patientunitstayid, observationoffset, value
        kernel: the self-defined kernel function for Gaussian Process Regressor

    Returns:
        data_full: the dataframe of aligned and interpolated data, including patientunitstayid, observationoffset, value
        data_full_index: the series of the index of the first occurrence of each patient
    """
```

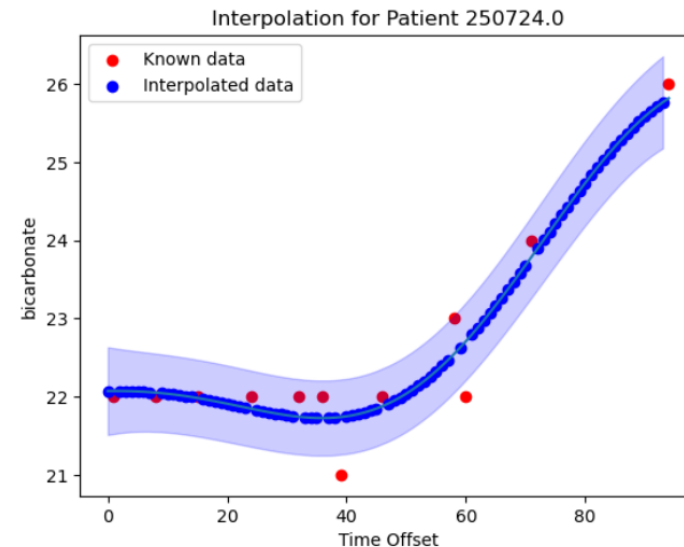
## 02 Aligning data



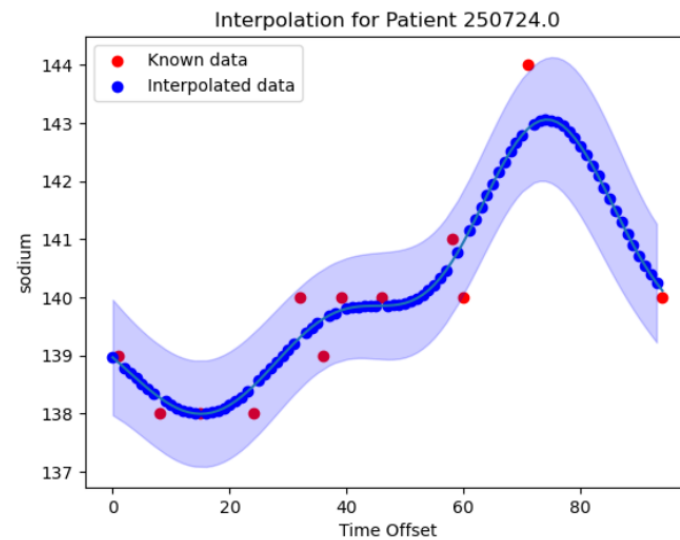
BUN



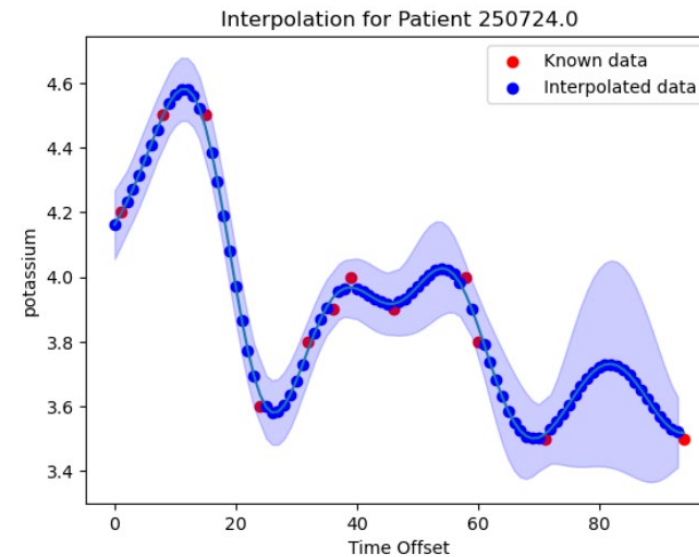
WBC



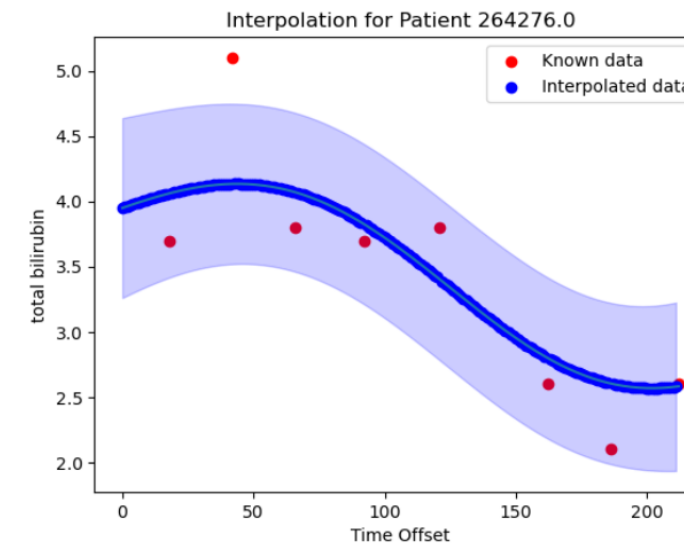
Bicarbonate



Sodium

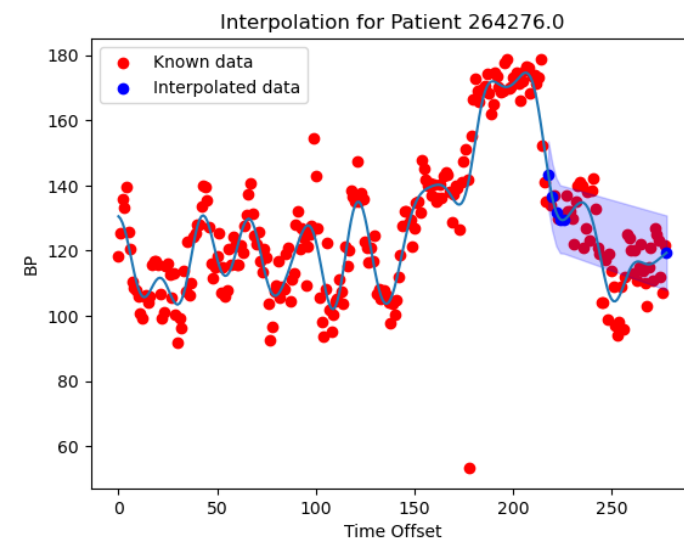
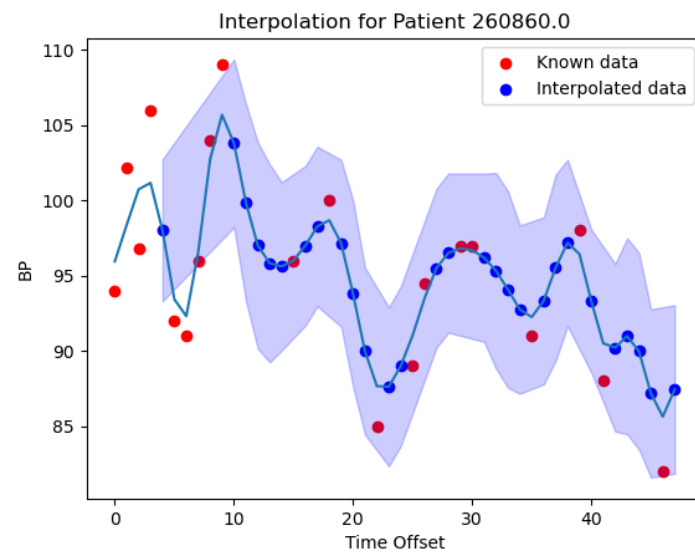
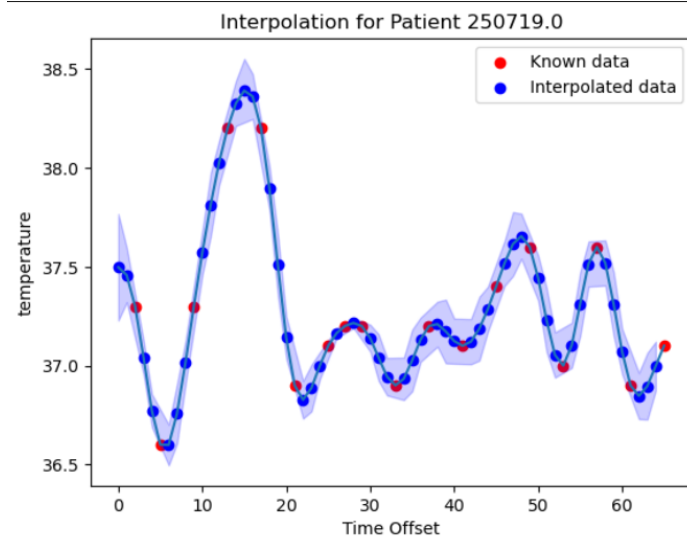
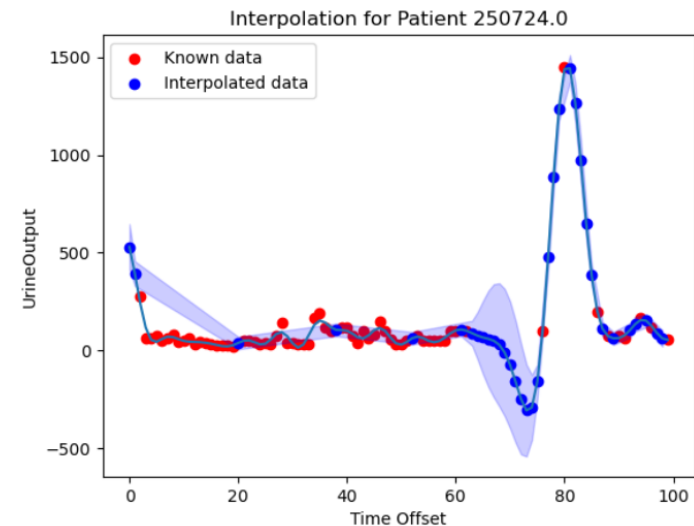
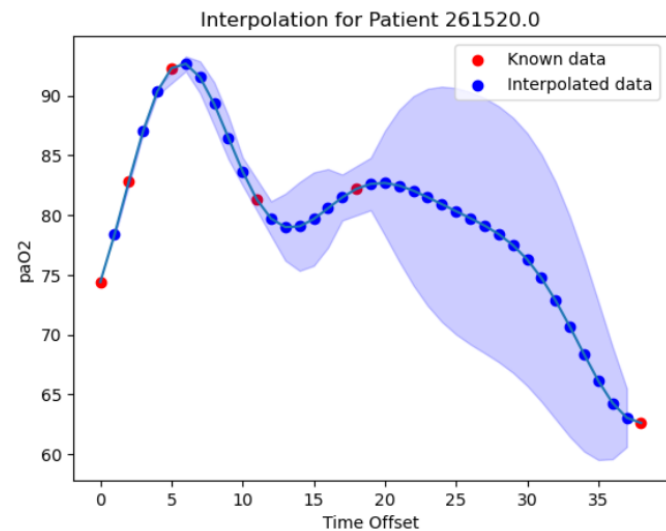
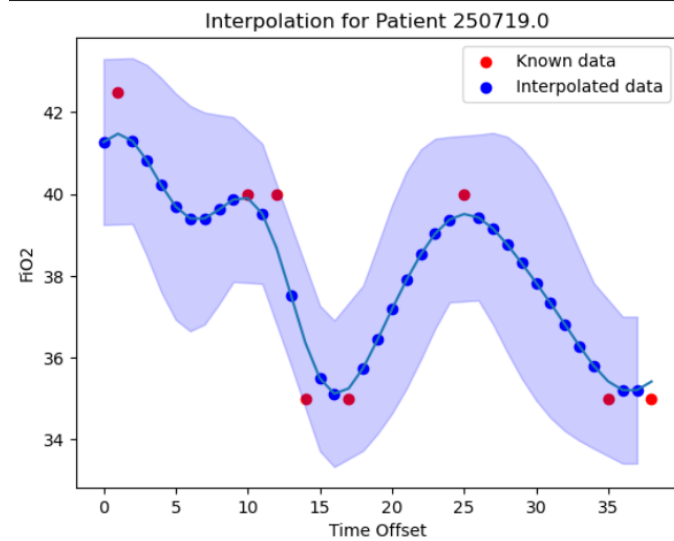


Potassium

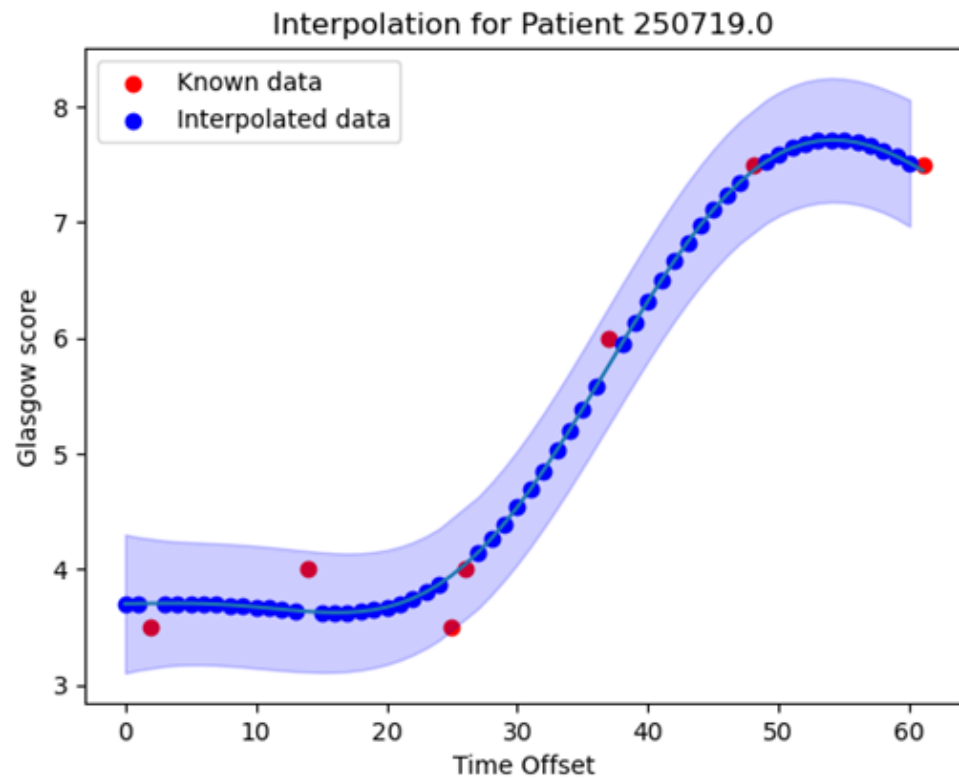


Bilirubin

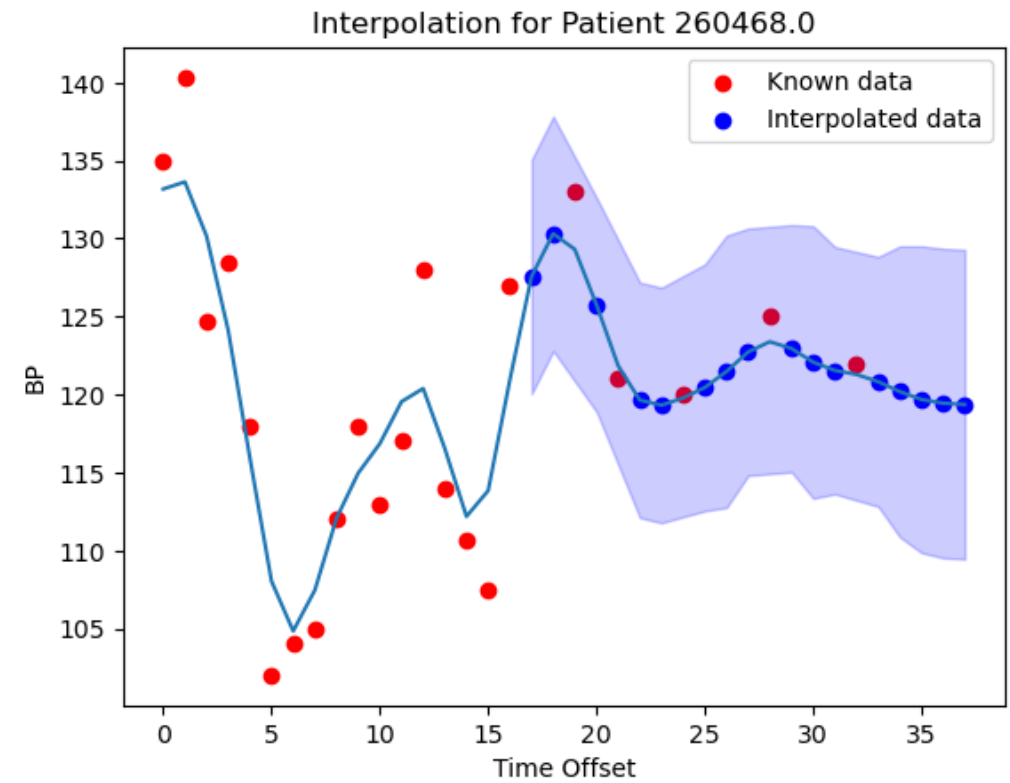
## 02 Aligning data



## 02 Aligning data

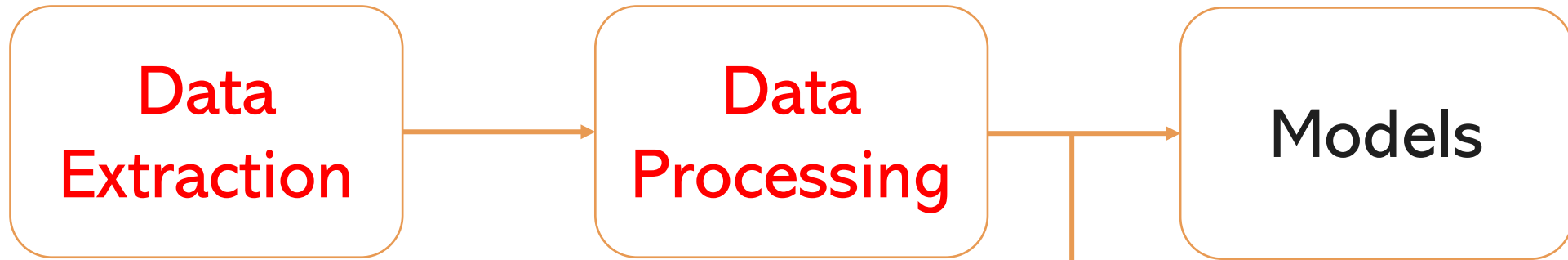


Glasgow



Heartrate

## 02 Our progress



Find out the **data interface** of the models

Process data from the **complete dataset** and save it in a **format callable by the model**



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