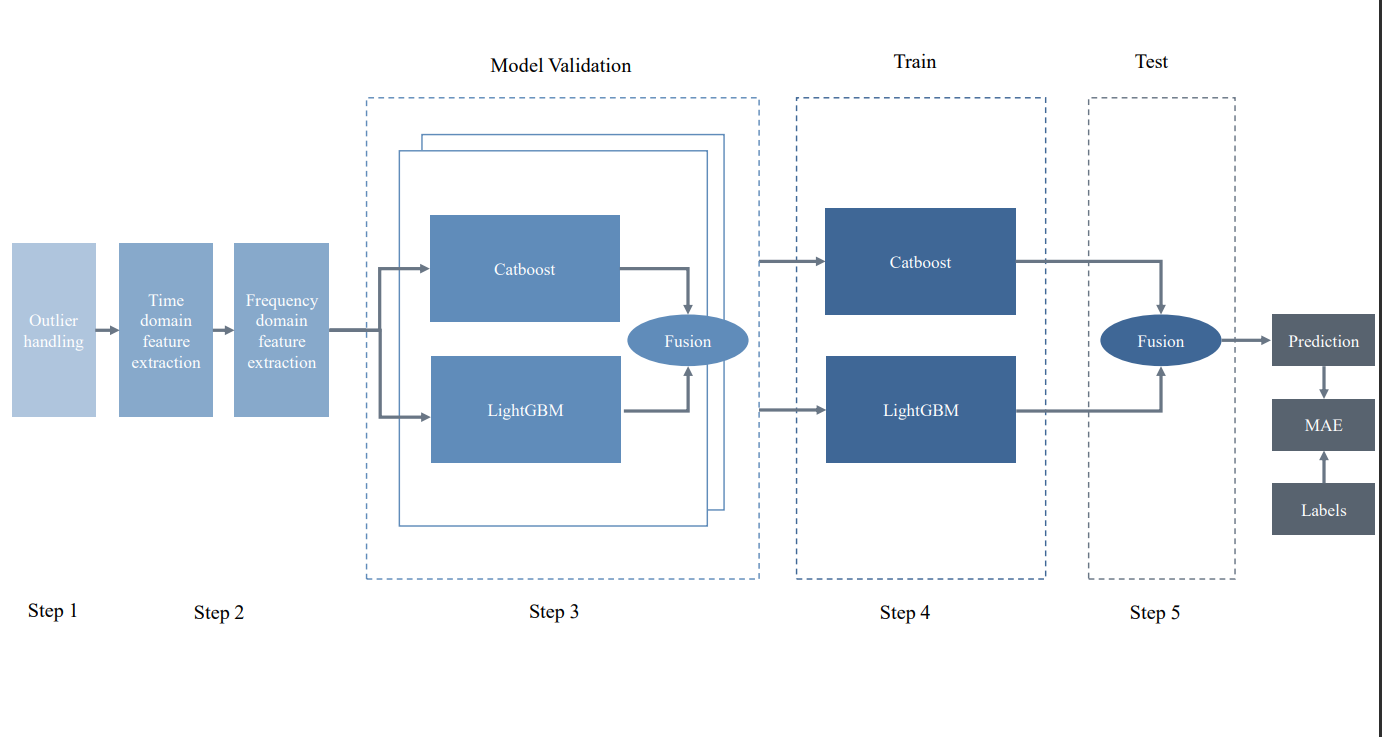
**Report**

**Jiayu Chen Hohai University**

Different algorithms were used for the prediction of the two indicators S and D, respectively. Indicator D uses a machine learning algorithm based on feature engineering. The indicator S, on the other hand, is difficult to mine to get the core features because of data bias, so a deep learning model is used.

**Predict D**



**Step1.** **Data pre-processing**

First, a high-pass filter is used to filter low-frequency signals below 1 Hz to remove noise.

Also, it is known from the physical meaning of the data that diastolic blood pressure cannot be greater than systolic blood pressure for the same person. Remove all data points where diastolic blood pressure is greater than systolic blood pressure.

The rest of the data are very clean, with no missing values and few outlier points. The distribution of the values to be predicted is normal and there are no anomalies such as long-tailed distributions.

**Step2. Extracting time-series features**

After feature screening, the valid time-series features are as follows: median, standard deviation, mean and kurtosis

**Step3. Extracting frequency domain features.**

Using Fourier variation, the signal is transformed to the frequency domain, and after feature filtering, the valid frequency domain features are as follows: main frequency bandwidth, peak, peak corresponding frequency, maximum frequency, minimum frequency, ECI, THD, mean, variance, kurtosis, and rate of change

**Step4. Model validation.**

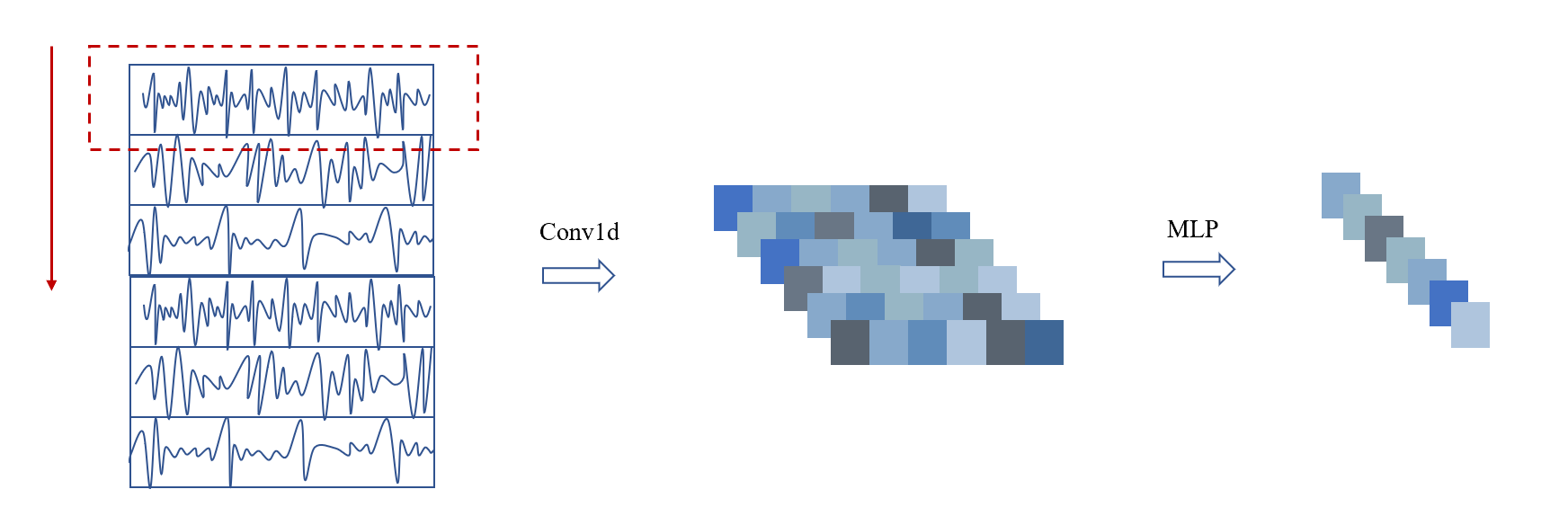
Xgboost, LightBGM, Catboost are three models that have performed well in data mining competitions. The Xgboost calculation requires a second-order derivative and is therefore not very suitable for optimization of MAE metrics.

Multi-fold validation of Catboost, LIGHTBGM and the integrated model of both are performed respectively. Model integration was found to perform well on 2, 3, and 4 fold validation.

**Step5. Model Training**

All training set data were used to train Catboost and LIGHTBGM and integrated.The equation for model integration is ，, are calculated in Step3.

**Predict S**



**Step1. Data pre-processing**

the original 10 s\*100 Hz signal is merged with the features mined when predicting the D-indicator.

**Step2. Building a Network**

Compression extraction of timing signals using 1-dimensional convolution. Next, dimensional compression is performed using multiple linear layers.