Time Series Pattern Recognition final report

36414067 Ryo Iida

2024/6/11

1 Visualization

1.1 TimeCourse plot

1.2 Detect Anomaly by looking

For both data, 図 1 and 図 2, the data around 217.5 can be seen as an anomaly.

2 Analysis

2.1 in time domain

For both data, 図 3 and 図 4, the data are periodic at approximately lag k = 210.

2.2 in frequency domain

For both data, 図 5 and 図 6, it can be seen that formants are formed at low frequencies.

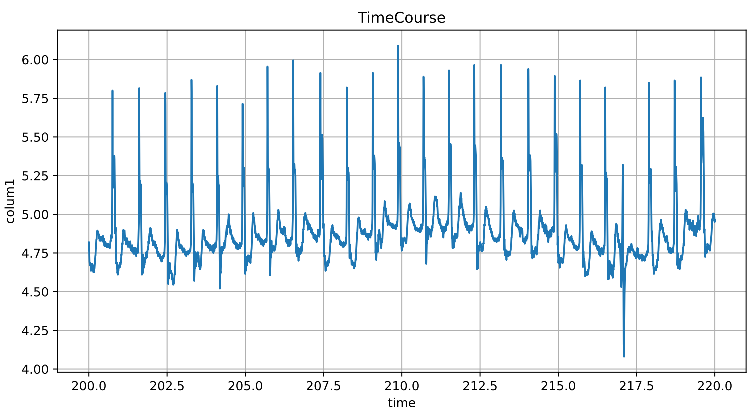

図 1: colum1: TimeCourse

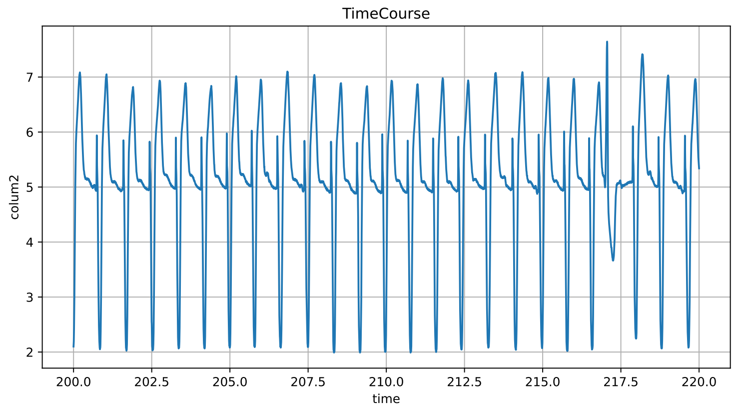

図 2: colum2: TimeCourse

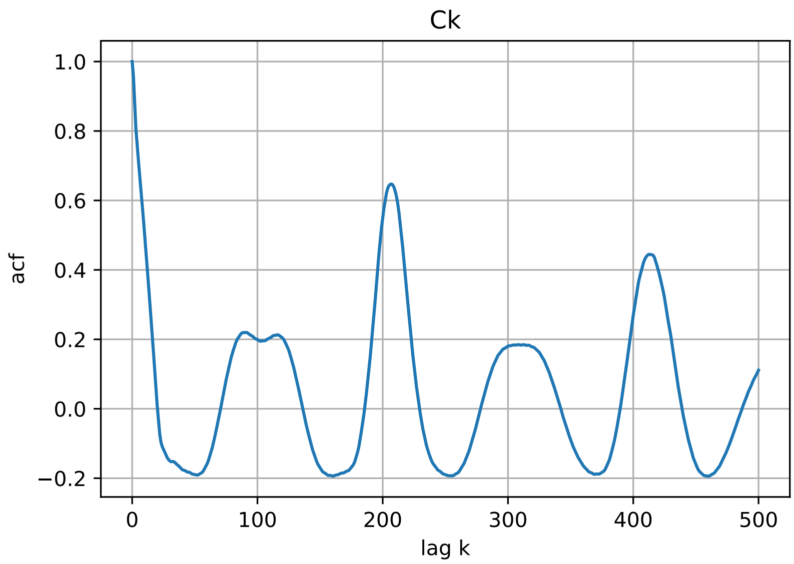

図 3: colum1: Auto covariance

3 Modeling

Keeping in mind that the data given from the analysis results so far is a periodic signal, the AR model is simple and has a low computational load, and is a basic model for capturing periodicity, while the NN model is a basic model for capturing periodicity. We use these two because we believe they have the ability to capture nonlinear patterns and are powerful when there is sufficient data. Regarding other models, model construction and parameter estimation may be more complex for SS models than for AR models, and RC models may be less effective depending on the amount of data and computational resources. Therefore, this time we chose the AR model and NN model because they are simple periodic signals.

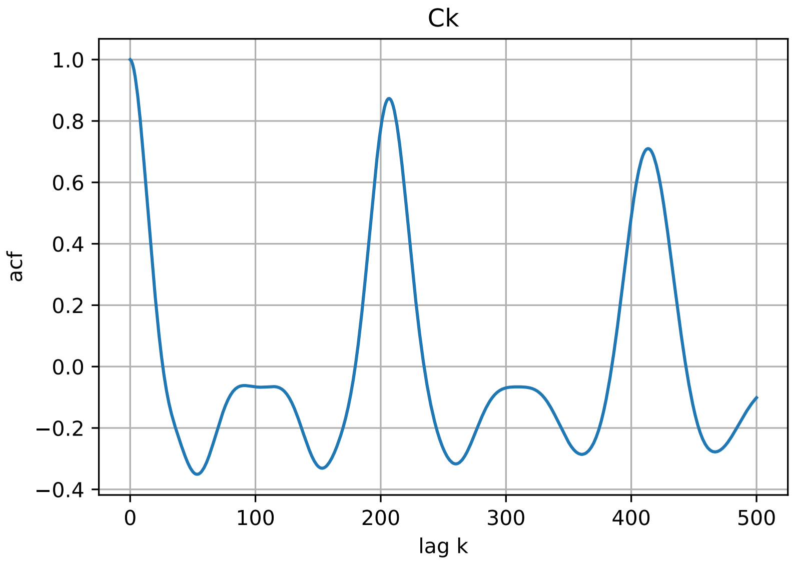

図 4: colum2: Auto covariance

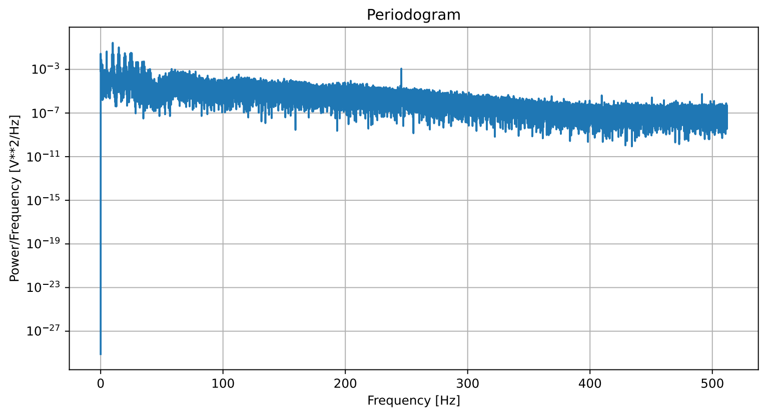

図 5: colum1: periodogram

4 Anomaly detection with One model

Looking at the two pieces of data, 図 7 and 図 8, it appears that the AR model continuously copies and pastes past periodic signals into future signals. However, because the predicted period gradually deviates, abnormality detection cannot be performed successfully, and the threshold value is exceeded even for parts that are not considered abnormal. For this reason, we considered the AR model to be too simple and unsuitable.

5 Anomaly detection with The other model

The error is the squared error as before. This time, we trained LSTM using three types of data: training data , validation data, and test data. However, it can be seen

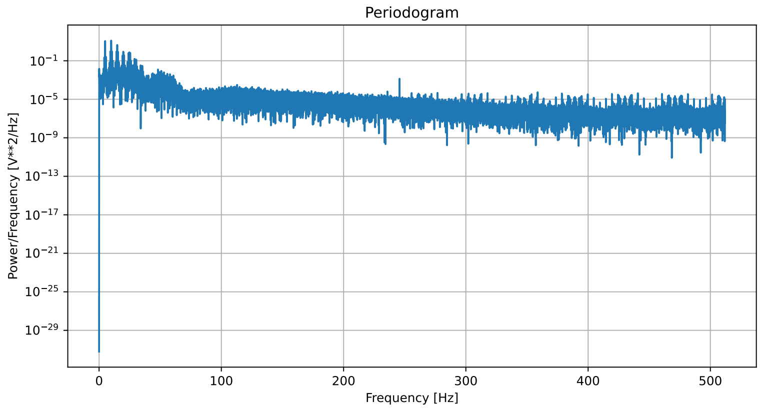

図 6: colum2: periodogram

that the fitting is correct even in areas that are considered to be anomaly detection. In column 2, when the threshold is set to 0.35, it can be recognized as above, but in column 1, an abnormality is detected at other times, so there is some concern. From this, considering that this anomaly could be predicted before the test data or verification data, we believe that this anomaly should be determined to be a normal value for this one learning data alone. We believe that new learning data separate from this will be required to determine that areas that are considered to be abnormal areas are abnormal.

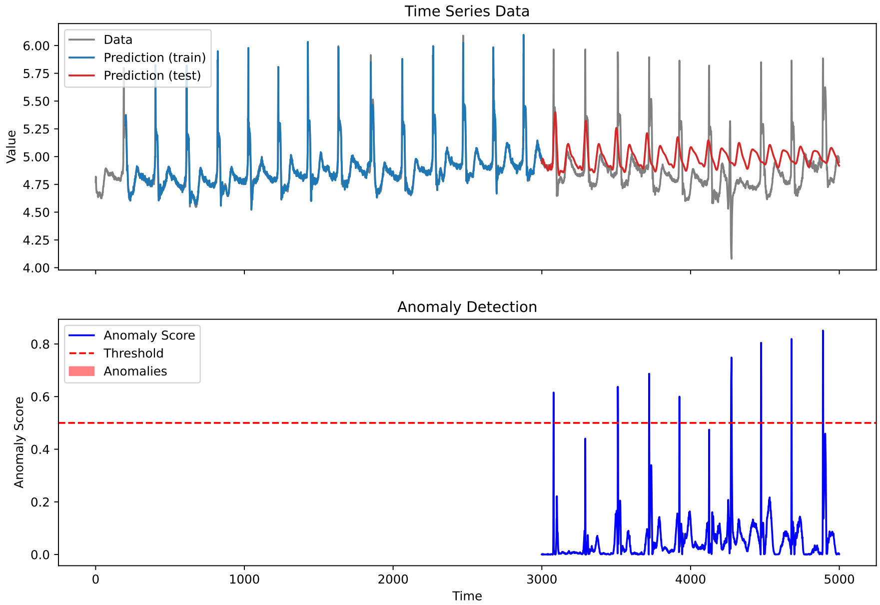

図 7: colum1: AR(200)model,threshold = 0.5

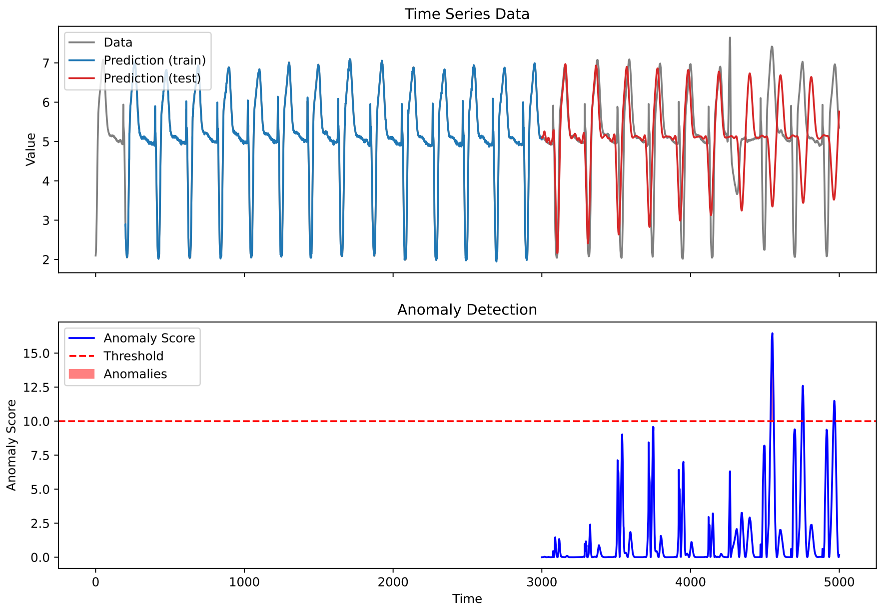

図 8: colum2: AR(200)model, threshold = 10

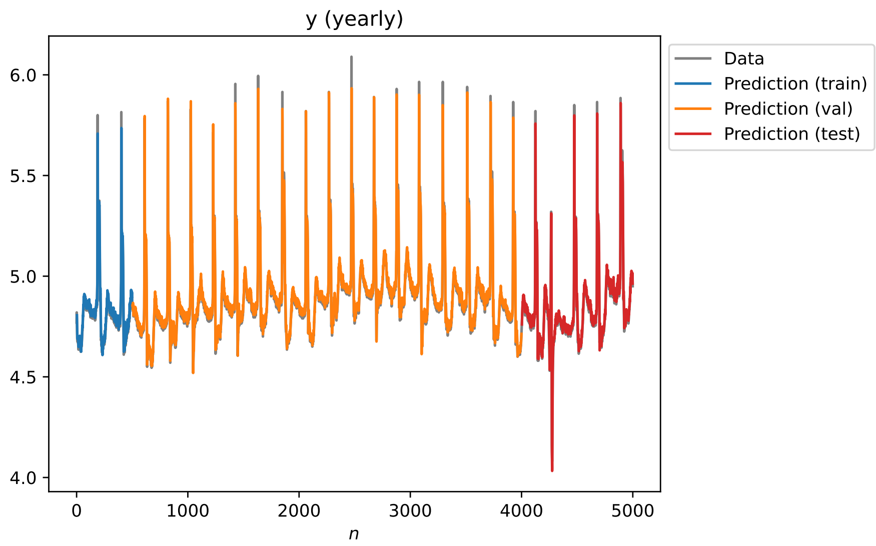

図 9: colum1: LSTM,train:0.1,validation:0.7,test:0.2

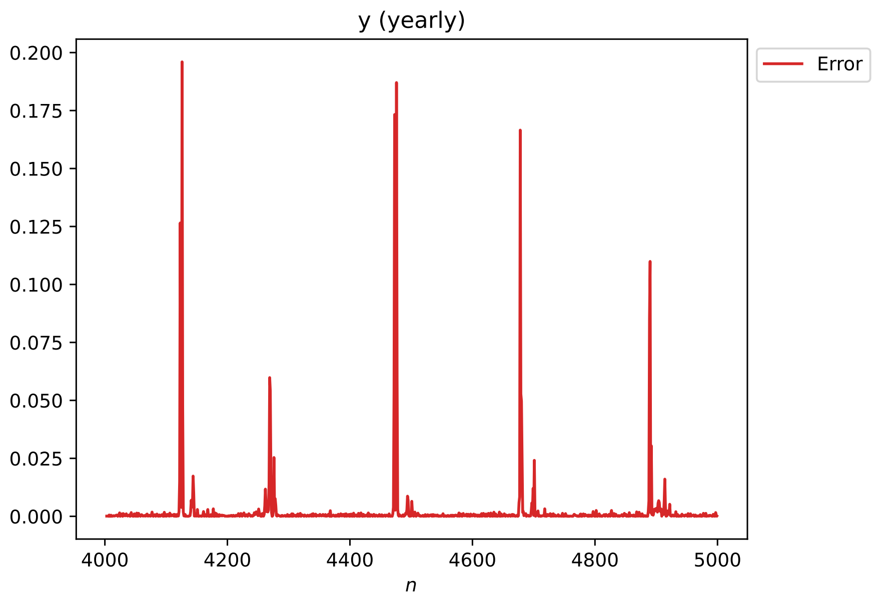

図 10: colum1: Error

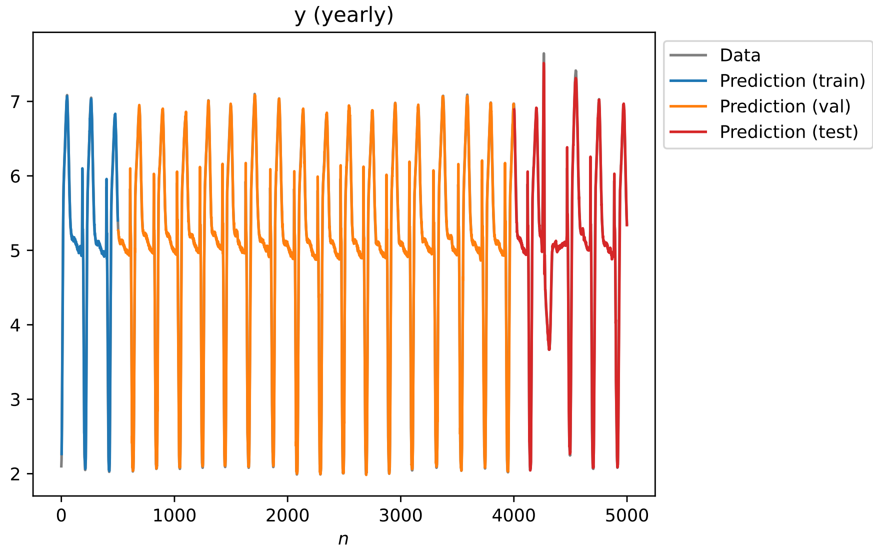

図 11: colum2: Error

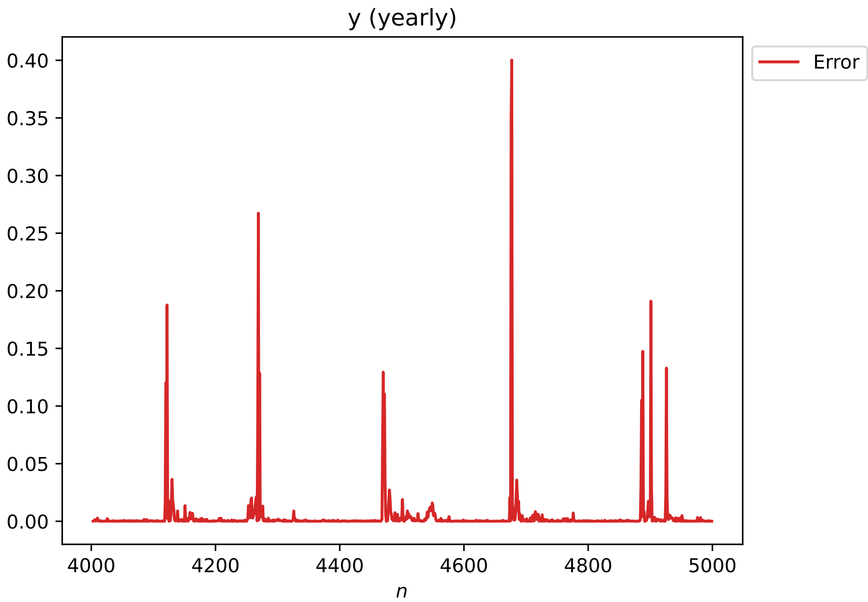

図 12: colum2: Error