1-17

Time Series Analysis by Data Augmentation Using Neural Network with Logistic Noise

†Ryosuke SHIMIZU †Tomiyuki FURUGAKI †Tomoya TAKATA
†Yoko UWATE †Yoshifumi NISHIO ‡Hidekuni TAKAO
†(Tokushima University) ‡(Kagawa University)

1. Introduction

The one of the most challenging problems about Neural Network (NN) is time series analysis. Therefore, it is important to provide positive feature to NN. In this study, chaostic noise is added to time series data because we estimate that chaostic noise can be used to increase the robustness of the network. By using computer simulations, we investigate how chaostic noise affects test accuracy of NN.

2. Proposed Method

In this study, chaotic noise caused by logistic map is used. Using this noise due to non-uniform output of this probability distribution increases the robustness of the network and improves the learning system expect. The logistic map behaves like various chaos depending on the parameters. The proposed method of this study is to generate noise with nonuniform frequency distribution and use this chaos. The formula of the logistic map is shown below. n means the number of steps. When X_n is determined, the next value X_{n+1} is uniquely determined.

$$X_{n+1} = aX_n(1 - X_n) \tag{1}$$

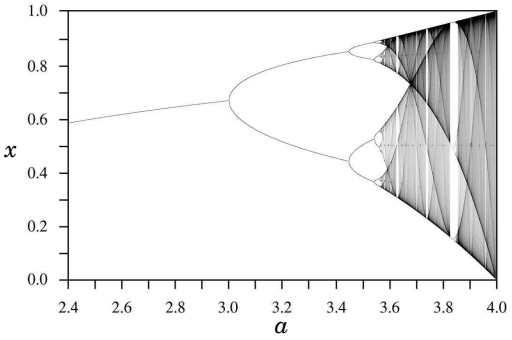

Figure 1: Bifurcation diagram of logistic map.

Figure 1 shows bifurcation diagram of logistic map when the parameter value is changed. The logistic map is in which the change in the value depends on what value is given to the parameter a, converging to a constant value, cyclically repeating multiple values vibrations, non-periodic and extremely complicated behavior called chaos. In this study, the values are set to $a = 4.0$ and $a = 3.828327$. When $a = 4.0$, it shows a behavior close to a random number called pure chaos. When we determine the parameter a=3.8238327, the logistic map shows the intermittency chaos.

3. Simulation Results

We investigate the relationship between the logistic noise in the data and the resulting test accuracy after training the NN with this data. In this study, three time series data are classified. Table 1 shows the number of the data with the 5000 points.

Table 1: The number of the time series data.

data	original	add noise
train data	336	672
test data	84	168

Table 2 and 3 show the test accuracy of the conventional method and the proposed methods using NN. The accuracy of each test was averaged 10 times. Also,by setting the learning rate to 0.001 and the epoch of the training to 200, the training accuracy has reached a maximum value 1.000. Table 3 shows the test accuracy with changing the parameter value a.

Table 2: Test accuracy of the conventional method.

	test accuracy
original	0.697

Table 3: Test accuracy of the proposed method.

	test accuracy
a=3.828327	0.887
a=4.0	0.889

Adding logistic noise to the time series data made it more accurate than the original. It can be said that the test accuracy of a=4.0 and a=3.828327 in Eq. (1) are better, compared to that of original data. It turned out that the test accuracy improves as a result of adding the logistic noise.

4. Conclusion

In this study, we proposed to add the logistic noise to improve the classification accuracy of time series data. In particular, pure chaos noise and intermittency chaos noise led to positive test accuracy results.

In the future, it is necessary to increase accuracy by leaving necessary noise and removing unnecessary noise and to add this noise to chaostic time series data.