Winter 12-6-2018

The application of Bounded Online Gradient Descent Algorithms for Kernel Based Online Learning in Tourist Number Forecasting

Qinghui Ren

Shenglin Li

Bo Song

Chen Chen

Ren, Q.H. Li, S.L. Song, B. & Chen, C. (2018). The application of bounded online gradient descent algorithms for kernel based online learning in tourist number forecasting. In *Proceedings of The 18th International Conference on Electronic Business* (pp. 800-804). ICEB, Guilin, China, December 2-6.

The application of Bounded Online Gradient Descent Algorithms for Kernel Based Online Learning in Tourist Number Forecasting

(Work in Progress)

Qinghui Ren, Army Logistics University of PLA, China, renqinghui007@163.com
Shenglin Li*, Army Logistics University of PLA, China, 1936423516@qq.com
Bo Song, Army Logistics University of PLA, China, 545641621@qq.com
Chen Chen, Army Logistics University of PLA, China, 433432@qq.com

ABSTRACT

With the upgrade of tourism informationization and the rise of intelligent city construction, the Internet of Things, Cloud Computing and other information technologies have been widely used in the tourism industry, making the development model and technical framework of intelligent tourism become a hot issue. Aiming at the problems of insufficient preparation and inadequate reception capacity of scenic spots in recent years, this paper proposes to apply machine learning algorithm to predict the number of tourists, so as to make an early response. In this paper, the characteristics of the application of the number of tourists are analyzed. The fixed buffer kernel online gradient descent algorithm is used to predict the number of tourists, and the actual number of tourists is brought into the algorithm for experiments. Finally, the rationality of the experimental process and results is analyzed.

Keywords: Smart tourism, online learning, kernel gradient decline, tourist number forecast;

*Corresponding author

INTRODUCTION

With the improvement of living conditions, the number of people who choose to travel has continued to rise in recent years, accompanied by a large number of tourists is the scenic spot far beyond the expected number of tourists prepared inadequate reception capacity. In this case, accurately predicting the number of tourists and making a good plan in advance according to the predicted results is the key to the above problems, and is also a key issue in the study of intelligent tourism. (Guan, 2016).

According to the modeling theory, there are two kinds of tourist quantity prediction: one is the traditional tourist quantity prediction model based on the statistical principle, which is based on the linear theory. But because the change of tourist quantity is a nonlinear dynamic system, it is affected by many factors, such as climate, advertisement and so on. As a result, the prediction accuracy is low and the application scope is limited (Kim, 2003; Huang, 2013; Ma, Liang & Lu, 2016). Another kind of machine learning prediction model is represented by neural network, which has good approximation ability to nonlinear, and the prediction accuracy of the number of tourists has been greatly improved (Deng & Lu, 2006; Dun, Zhang, Liu & Zhang, 2017; Chang & Tsai, 2017; Li & Zhang, 2010; Gjylapi, Durmishi & Musaraj, 2014). However, the traditional neural network training process needs to modify the weights of the network repeatedly, which requires a large amount of calculation, and there are some problems such as over-fitting and dimension disaster (Huang, Zhu & Siew, 2006). Extreme Learning Machine (ELM) is a new kind of feed forward neural network. According to Moore-Penrose generalized inverse matrix theory, the training iteration is transformed into solving linear equations, and the network training can be completed at one time. Compared with the traditional neural network, it needs many iterations to determine the network output weights. ELM greatly improves the network prediction modeling efficiency (Xia, Zhang, Weng & Ye, 2012).

However, most of the previous methods used fixed data to establish the tourist number prediction model, which belongs to the typical off-line prediction model. However, the number of tourists has the characteristics of time series, real-time change, stochastic and strong time-varying. It is difficult to predict the number of tourists accurately by using offline model (Feng, Huang, Lin & Gay, 2009). Combining with the time series characteristics of the number of tourists, this paper adopts the Bounded online gradient descent algorithm for scalable kernel based online learning (BOGD) to predict the number of tourists, and uses the actual number of tourists to calculate and predict, according to the experimental results. The effectiveness and defects of BOGD algorithm in the application of tourist volume prediction are discussed.

BOGD ALGORITHM

BOGD algorithm is an on-line algorithm proposed by Zhao, Wang, Wu, Jin, and Hoi (2012). Its kernel function is used to realize nonlinear classification and regression. Gradient descent method is used to optimize the classifier. The core idea of the algorithm is as follows:

Let the collected time series data be $\{(x_t, y_t), t \in [T]\}$, among them $x_t \in R^d$, $y_t \in R$, $[T] = \{1, \dots, T\}$.

The online learning model of BOGD algorithm is:

$$L(f) = min_{f \in H} l(f(x_t), y_t) + \frac{\gamma}{2} \|f\|_H^2 \tag{1}$$

γ is the regularization parameter. When a new set of samples (x_t, y_t) arrives, the online gradient descent method is used to update the classifier, and the updating rules are as follows:

$$f_{t+1} = f_t - \theta \partial_f L(f) | f = f_t \tag{2}$$

θ is the learning rate constant, which can also be understood as falling step size. According to the regeneration characteristic of kernel function $K(\cdot, \cdot)$:

$$\langle f, K(x, \cdot) \rangle_H = f(x), for\ x \in X \tag{3}$$

And

$$\partial_f l(y_t f(x_t)) = y_t l'(y_t f(x_t)) K(x_t, \cdot) \tag{4}$$

Bringing (4) into the update expression (2) can be obtained:

$$f_{t+1} = (1 - \theta\gamma) f_t - \theta y_t l'(y_t f(x_t)) K(x_t, \cdot) \tag{5}$$

Let $f_1 = 0$, f_t can be written as a form of kernel expansion:

$$f_t(x) = \sum_{i=1}^{t-1} \alpha_i y_i K(x_i, x), x \in X \tag{6}$$

The coefficient of step *t* is updated to:

$$\alpha_t = \theta y_t l'(y_t f_t(x_t)), \quad i = t; \tag{7}$$

$$\alpha_i = (1 - \theta\gamma) \alpha_i, \quad i < t; \tag{8}$$

In the above updating process, with the deepening of the calculation, support vector will be generated continuously. If the amount of data is infinite, then an infinite number of support vectors will be generated. The calculation of support vector weight α_i will occupy a large amount of memory and computing resources. To solve this problem, the number of support vectors N_{SV} is limited in the BOGD algorithm. The upper limit of the number of support vectors is B, which is called intercept parameter. When $N_{SV} < B$ is satisfied in the above calculation process, iterative calculation is carried out according to the above steps. As each iteration process, the coefficient $\alpha_i (i \neq t)$ will be reduced by $(1 - \theta\gamma)$. After B times iteration, it will be reduced to $(1 - \theta\gamma)^B \alpha_i$. Therefore, when $N_{SV} > B$ is used, the support vector with small coefficients can be discarded, which only leads to very small errors.

The input and initialization of the BOGD algorithm are as follows:

Input: sample sequence is $(x_t, y_t)_{i \in N} \in (X \times Y)^\infty$, regularization parameter is $\gamma > 0$, intercept parameters satisfies $B \in N$, learning rate is $\theta \in (0, \frac{1}{\gamma})$, loss function is $l: R \times Y \to R, H$, space reproducing kernel is $K(\cdot, \cdot)$;
Initialize: number of support vectors is $S_1 = 0$, $f_1 = 0$;

The calculation process of the BOGD algorithm is as shown in Figure 1.

THE APPLICATION OF BOGD ALGORITHM TO PREDICT THE NUMBER OF TOURISTS

The number of tourists is a typical unstable time series forecasting problem. Although the number of tourists changes from time to time, it has obvious time series. Therefore, BOGD algorithm is used to predict the number of tourists. The online characteristics of the algorithm can meet the requirements of real-time processing of the number of tourists. At the same time, the fixed buffer kernel gradient descent method ensures that the calculation process will not take up unlimited resources, which guarantees the efficiency of the algorithm.

Composition of Time Series

The training data in this paper are based on the original number of tourists and the number of tourists in Jiuzhaigou from 06/01/2012 to 06/01/2017.

Data details are shown in Table 1.

for $t = 0,1,\ldots T$ do

Receive x_t;

Predict $\hat{y}_t = f_t(x_t)$;

Receive y_t and suffer loss $l(y_t f(x_t))$;

If $l(y_t f(x_t)) = 0$ then

$f_{t+1} = (1-\eta\lambda) f_t$ and $S_{t+1} = S_t$;

Else

If $|S_t| < B$ then

$f_{t+1} = (1-\eta\lambda) f_t - \eta y_t l'(y_t f(x_t)) K(x_t, \bullet)$

And $S_{t+1} = S_t \cup \{t\}$

Else

$f_{t+1} = \sum (1-\eta\lambda)\alpha_i y_i K(x_t, \bullet) - \eta y_t l'(y_t f(x_t)) K(x, \bullet)$

$S_{t+1} = S_t \cup \{t\} / \{1\}$

End if

End if

End for

Figure 1: The Calculation Process of BOGD Algorithm

Table 1 Details of Tourist Quantity for Training

Spot name	Start time	End time	Data quantity	Owed Country
Jiuzhaigou	06/01/2012	06/01/2017	1827	China

As shown in Table 1, the experimental data are from Jiuzhaigou Scenic Area in China, with a total of 11 data items. The online access address of the data is https://www.jiuzhai.com/news/number-of-tourists.

Data Preprocessing

In order to improve the training efficiency and the generalization ability of calculation, it is necessary to compress the range of actual input data. For this reason, we normalized the original data of the number of tourists.

Let x_{ti} be the i th component of the data arriving at t time. The maximum value of all data is MAX and the minimum value is MIN. The data are normalized by Equation (9):

$$x'_{ti} = \frac{x_{ti} - MIN}{MAX - MIN} \tag{9}$$

The normalized data are brought into the algorithm for calculation.

Forecast of Tourists and Result Analysis

The normalized data are brought into the algorithm for calculation. The actual and predicted numbers of tourists for the day are shown in Figure 2.

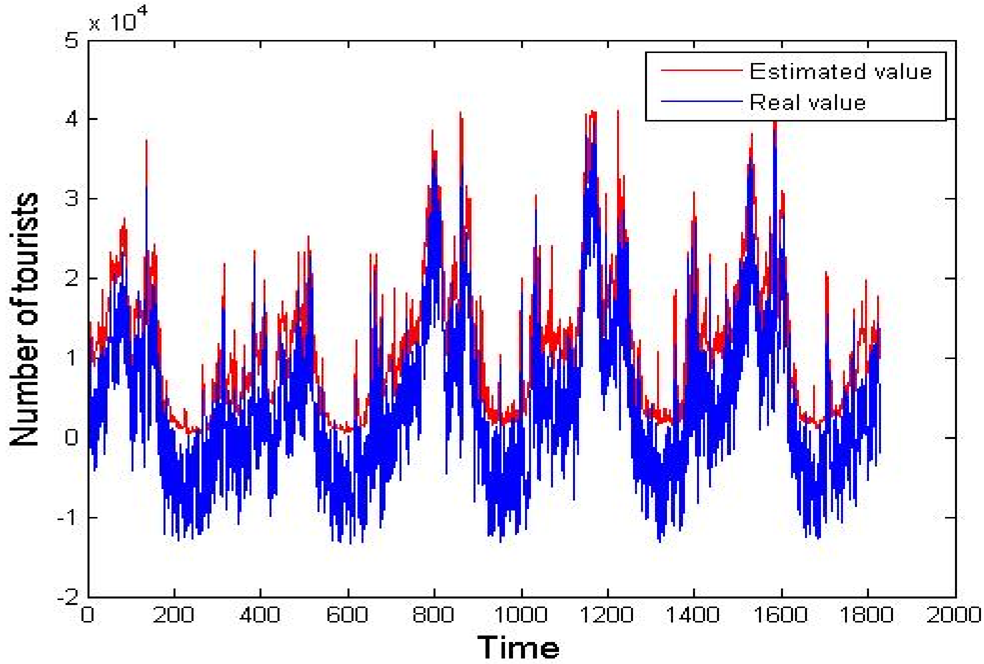

Figure 2: Comparison of Actual and Predicted Tourist Numbers

As can be seen from the above figure, the number of tourists in this scenic spot shows obvious fluctuations. This is due to the obvious increase in the number of tourists in holidays, while winter is the off-season for tourism. A careful observation of the actual and predicted number of tourists reveals that the predicted overall trend is in good agreement with the actual value by using BOGD algorithm to calculate the number of tourists. However, in the single-day data, the accuracy of BOGD algorithm in predicting the number of tourists is not high.

CONCLUSION

This paper proposes to apply BOGD algorithm to study and predict the number of tourists. From the experimental results, it can be seen that the overall trend of the predicted data is similar to the actual data, so it can be used to predict the general trend of the number of tourists. However, it can be seen from the enlarged detail chart that the predicted data value of each day is still quite different in accuracy from the actual data.

After analysis, this result may occur for the following reasons. The BOGD algorithm used in this paper needs to preprocess the data and normalize all data to [0, 1]. However, for the online algorithm, we do not know the sample population before the regression calculation. Therefore, this method of calculating the maximum and minimum values of samples in the normalization process is not appropriate. In addition, the number of tourist changes in real time, and the accuracy of the experimental design method of predicting the number of tourists in days may not be guaranteed.

Although the above problems still exist in the experiment itself, this research process is a preliminary attempt to predict the number of tourists using the classic online algorithm and lays a good foundation for future research.

ACKNOWLEDGMENT

This work is supported by the Department of Military Logistics, Army Logistics University of PLA.

REFERENCES

[1] Guan X L. (2016). Research of problems existing in the tourism public service and countermeasure based on the customer satisfaction measurement.
[2] Kim K J. (2003). Financial time series forecasting using support vector machines. Neurocomputing, 55(1): 307-319.
[3] Huang H C. (2013). The construction of a Visitor number forecasting model for the Shihmen Reservoir recreation area using artificial intelligence technology. Journal of Theoretical & Applied Information Technology.
[4] Ma Y C, Liang C Y, Lu C H. (2016). The Weekly Visitor Numbers Forecasting of Tourist Off-season in a Scenic Spot Advanced Science and Technology. 198-204.
[5] Deng X T, Lu Y Q. (2006). The Application of BP Neural Network in Forecasting China's Inbound Tourist Population. Tourism Science, 20(4): 49-53.
[6] Sun, Y., Zhang, H. L., Liu, P. X., & Zhang, J. (2017). Forecast of tourism flow volume of tourist attraction based on degree of tourist attention of travel network: a case study of baidu index of different clients. Human Geography.
[7] Chang Y W, Tsai C Y. (2017). Apply Deep Learning Neural Network to Forecast Number of Tourists. International Conference on Advanced Information NETWORKING and Applications Workshops. IEEE, 259-264.

[8] Li J B, Zhang H Q. (2010). Application of BP neural network model on forecasting the number of tourists in Hainan province. International Conference on Computer and Communication Technologies in Agriculture Engineering. IEEE, 502-507.

[9] Gjylapi D, Durmishi V, Musaraj A. (2014). Artificial neural networks in forecasting tourists' flow, an intelligent technique to help the economic development of tourism in Albania. Academicus--International Scientific Journal, MMXIV(10): 202-211.

[10] Huang G B, Zhu Q Y, Siew C K. (2006). Extreme learning machine：Theory and applications. Neurocomputing, 70(1/3): 489-501.

[11] Xia, M., Zhang, Y., Weng, L., & Ye, X. (2012). Fashion retailing forecasting based on extreme learning machine with adaptive metrics of inputs. Knowledge-Based Systems, 36, 253-259.

[12] Feng, G., Huang, G. B., Lin, Q., & Gay, R. K. L. (2009). Error minimized extreme learning machine with growth of hidden nodes and incremental learning. *IEEE Trans. Neural Networks*, *20*(8), 1352-1357.

[13] Zhao, P., Wang, J., Wu, P., Jin, R., & Hoi, S. C. (2012). Fast bounded online gradient descent algorithms for scalable kernel-based online learning. International Conference on Machine Learning. *arXiv preprint arXiv:1206.4633*.