병원 외래환자수의 예측을 위한 시계열 데이터처리 딥러닝 시스템

조준모*

Time Series Data Processing Deep Learning system for Prediction of Hospital Outpatient Number

Jun-Mo Jo*

요 약

딥러닝 기술의 도래로 인하여 수많은 산업과 일반적인 응용에 적용됨으로써 우리의 생활에 큰 영향을 발휘하고 있다. 특정한 분야의 문제를 해결하기 위해서는 그 문제에 적합한 딥러닝 모델을 작성해야 한다. 근래에는 COVID-19 사태로 인하여 다양한 문제들을 딥러닝으로 해결하고자 하는 사례들이 늘고 있다. 이러한 일환으로 본 논문에서는 갑자기 급증할 수 있는 병원의 외래환자들을 미리 예측을 위한 시계열의 딥러닝 모델을 제시하고자 한다. 제시하는 딥러닝 모델은 주피터 노트북에서 케라스로 작성하였다. 예측결과는 실제 데이터와 그래프로 비교하며 유효성 데이터를 활용하여 과소적합과 과대적합의 여부를 손실률로 분석할 수 있도록 하였다.

ABSTRACT

The advent of the Deep Learning has applied to many industrial and general applications having an impact on our lives these days. Certain type of machine learning model is needed to be designed for a specific problem of field. Recently, there are many instances to solve the various COVID-19 related problems using deep learning model. Therefore, in this paper, a deep learning model for predicting number of outpatients of a hospital in advance is suggested. The suggested deep learning model is designed by using the Keras in Jupyter Notebook. The prediction result is being analyzed with the real data in graph, as well as the loss rate with some validation data to verify either for the underfitting or the overfitting.

키워드

Ⅰ. Introduction

The advent of the Deep Learning has applied to many industrial and general applications having an impact on our lives these days. Certain type of machine learning model is needed to be designed

for a specific problem of field. Deep learning has made possible the concept of self-driving cars, automatic intelligent web search, user based speech recognition software, personalized marketing and so on. In recent years the management of patient flow is one of the main challenges faced by many hospital establishments, in particular emergency departments. So the world wide research related in this field is underway in many universities, companies and research facilities. The researches are related to the supervised and unsupervised learning methods as well as the field of the deep learning. Such observation can define a new class or assign a new class to an existing class. The supervised and unsupervised are investigated its properties in the classification of post graduate students according to their performance during the admission period[1].

Introduction of cognitive reasoning into a conventional computer can solve problems by example mapping like pattern recognition, classification and forecasting. Artificial Neural Networks provides these types of models. These are essentially mathematical models describing a function. And the application function as a classification model. Finally, a learning rule such as supervised, unsupervised, and the reinforcement methods[2-3].

This paper is on the predicting the numbers of hospital patient using deep learning models for time series data processing. So, in section II, a time series data preparation as preprocessing for the deep learning is explained. In section III, the deep learning models such as LSTM and GRU for the time series data processing applied in this paper is elaborated. Then in section IV, the result of the experiment is analyzed and there is conclusion in the section V.

II. Time-Series Data Preprocessing

The research on a time-series data as an input of the deep learning are emerging these days. The time series input features needs to be utilized in a limited prediction timing. And, there is a research on the case of corporate bond credit rating forecast, corporate sample is limited because only large companies are selected for corporate bond credit rating. To address limitations of prior research, this study attempts to implement a predictive model with more sample companies, which can adjust the forecasting point at the present time by using the credit score information and corporate information in time series[4]. So the preproccesing process should be considered the neural models such as the RNN(Recurrent neural network) based stateful LSTM(Long Short Term Memory) model. The example of the LSTM model is shown as Fig. 1[5].
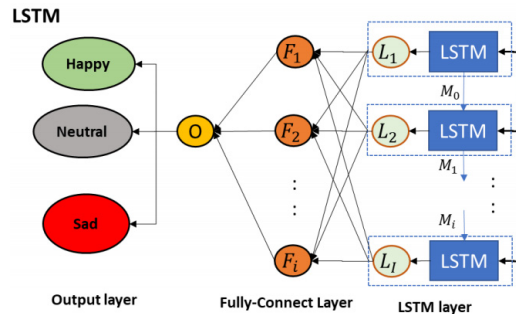

Fig. 1 Example of LSTM model

Machine Learning methods requires the fine tuning of the parameters and also feasible number of the data set. Therefore, choosing the best performance of the learning algorithm is important in the real world. The best performance is decided by not only the recognition ratio but also the time of the simulation. And also for a particular data set does not guarantee the precision and accuracy for another set of data whose attributes are logically different from the other. However, the conditions of

a particular method can significantly outperform others on a given application problem. Meta learning is moving in this direction, trying to find functions that map data sets to algorithm performance. After a better understanding of the strengths and limitations of each method, the possibility of integrating two or more algorithms together to solve a problem should be investigated. The objective is to utilize the strengths of one method to complement the weaknesses of another. If we are only interested in the best possible classification accuracy, it might be difficult or impossible to find a single classifier that performs as well as a good ensemble of classifiers[6].

The normalization method using in this experiment is the Min-Max method. the data is scaled to a fixed range, usually 0 to 1, the cost of having this bounded range is that it will end up with smaller standard deviations. A Min-Max scaling equation is as follows:

$$X_{norm} = \frac{X_{old} - X_{\min}}{X_{\max} - X_{\min}}$$

III. Time-Series Deep Learning Model

In order to predict the numbers of the hospital's patients, the recursive process is need in the neural network as shown in Fig. 2 such as RNN[7].
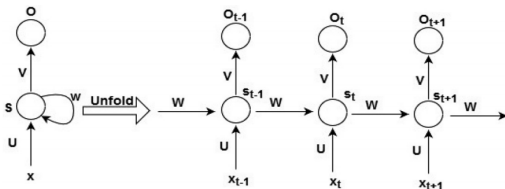

Fig. 2 Basic RNN MODEL

The latest models of RNNs such as LSTM and GRU are elaborated next.

3.1 LSTM(Long Short Term Memory)

A LSTM network can be applied to time series forecasting. There are many types of LSTM models that can be used for each specific type of time series forecasting problem.

Sigmoid layer in LSTM enables the network to have control over the flow of information. When the output of the sigmoid layer, as the following shown below, is zero, no information is going through. On the contrary, when the value is equal to one the system lets everything goes through[7].

$$sigm(x) = \frac{1}{1 + e^{-x}}$$

Due to the important role played by emotion in human interaction, affective computing is dedicated in trying to understand and regulate emotion through human-aware artificial intelligence as well[8, 9]. Moreover, the series of data like the numbers of patients visiting hospital on a regular basis can be adapted to the deep learning to predict the numbers in the future due to the COVID-19 Pandemic. In order to apply the application, the proper neural network model is required to accept a proper series of data.

3.2 GRU(Gated Recurrent Unit)

Gated Recurrent Units(GRUs) GRUs are improved version of standard recurrent neural network. And it is a gating mechanism in recurrent neural networks. To solve the vanishing gradient problem of a standard RNN, GRU uses, so called, update gate and reset gate. So, the structure of the GRU is similar with the LSTM with a forget gate, but has fewer parameters than LSTM without an output gate.

The performance of the GRU on certain tasks of polyphonic music modeling, speech signal modeling

and natural language processing was found to be similar to that of LSTM. They can be trained to keep information from long ago, without washing it through time or remove information which is irrelevant to the prediction. The GRU has been shown to exhibit better performance on certain smaller and less frequent data sets.

3.3 Predicting Numbers of Patients

The hospital data as an input of a neural network is obtained in the Internet and the data contains the periotic numbers of hospitals patients in England. The data is provided from the NHS England, and it is published in 2019. It is the monthly activity data relating to elective and non-elective impatient admissions(FFCEs) and outpatient referrals and attendances for first consultant outpatient appointments.

In order to obtain the future numbers of patients, the original data is divided into training data and test data. Since the data is in the time periotic order, the first 90% of the data is used for the training, and the rest of 10% data is compared with the predicted result for the evaluation. The final result is also considered to avoid overfitting and underfitting with validation data.

The neural network is coded in Keras for the experiment. It is a powerful and easy to use free open source Python library for developing and evaluating deep learning models. It wraps the efficient numerical computation libraries of Tensorflow and allows you to define and train neural network models in just a few lines of code.

IV. Result and Analysis

For the neural network the Tensorflow in Keras is used shown as Table 1. For the better performance the Dropout and Flatten method is applied to the network.

Table 1. Tensorflow in Keras

```
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, Dropout
from tensorflow.keras.layers import Flatten, LSTM
```

The three neural network is trained such as fully connected, LSTM, and GRU. The LSTM model is designed as follows shown in Table 2. Every node of the neural network is fed into the preceding data recursively, so the 'return_sequences' parameter is set to 'True'. And the 'dropouts' is used for randomly dropping out nodes during training.

Table 2. LSTM Neural Network Model

```
model.add(LSTM(32, return_sequences=True,
          input_shape=(1, lookback)))
model.add(LSTM(16))
model.add(Dropout(0.5))
model.add(Dense(1))
```

There are not much difference in the prediction result among them since the neural network is adjusted with not only the Dropout and Flatten, but also the types of the loss function and the optimization function for the validation result.

However, the fully connected model showed high loss at the beginning unlike the others as shown Fig. 3 and Fig. 4.
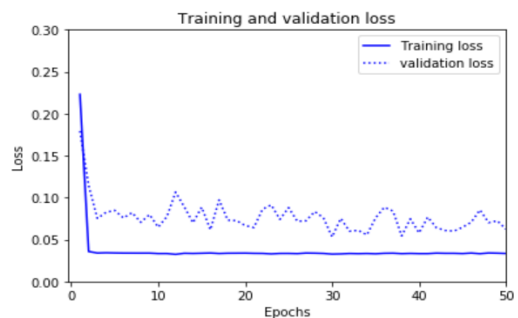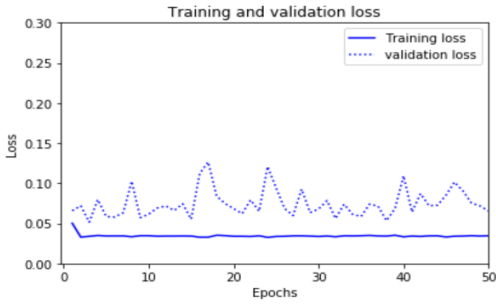

Fig. 3 Loss of Fully Connected Model

Fig. 4 Loss of GRU Model

The graph of entire data and the predicted data is shown as Fig. 5. The comparison of the test and the entire data graph shows that they are not perfectly matches but the trend is similar. The result of other models shows the similar result.
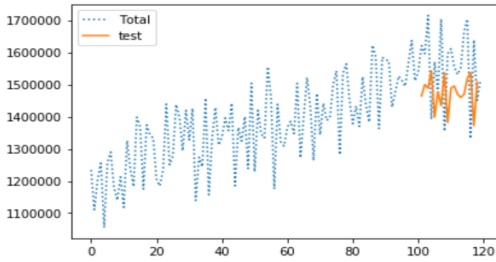

Fig. 5 Predicted data of GRU Model

The following graph Fig. 6 shows the only predicted area of the patients numbers which indicates the real patients numbers of 1,550,000 compare to the predicted numbers of 1,470,000 patients at the beginning.
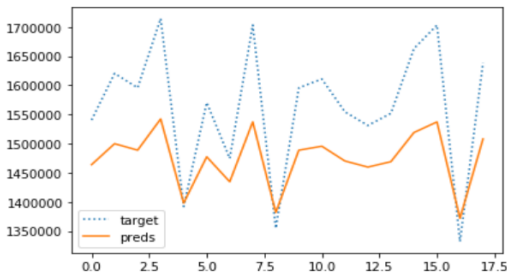

Fig. 6 Comparison of test predicted and targets

The entire(Total) target data and the entire predicted data are compared in Fig. 7. The 'Predict' data is produced with the preceding train data. In other words, the inputs are fed into the next gate to obtain for the next predict data recursively. So the result looks like the Fig. 6.
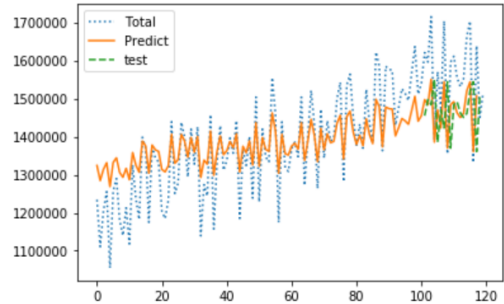

Fig. 7 Entire predicted and targets

The prediction numbers are not perfectly matches but it is acceptable numbers to figure out the trend of emerging or decreasing.

V. Conclusion

The prediction of the numbers of hospital patient using deep learning models is designed, executed, and analyzed in Keras. Unlike the image related neural network, to predict the numbers of hospital patients needs recursive types of neural network. So, the preprocessing of the data for the deep learning model is preceded. The three neural network models such as the fully connected, the LSTM, and the GRU models are experimented and

317

compared with the results. First of all, the results are similar among them in the prediction numbers as well as the loss rate. However, the prediction numbers are not perfectly matches but it is acceptable numbers to figure out the trend of emerging or decreasing. So for the future research, the suggested model can be applied to the patients of hospitals in Korea.

Reference

[1] R. Sathya and A. Annamma, "Comparison of Supervised and Unsupervised Learning Algorithms for Pattern Classification," *IJARAI,* vol. 2, no. 2, 2013, pp. 34-38.

[2] R. Sathya and A. Abraham, "Unsupervised Control Paradigm for Performance Evaluation," *International Journal of Computer Application*, vol. 44, no. 20, 2012, pp. 27-31.

[3] X. C. Yin, X. Yin, K. Huang, and H. W. Hao, "Robust text detection in natural scene images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 36, no. 5, 2014, pp. 970-983.

[4] H. Lee and S. Oh, "LSTM-based Deep Learning for Time Series Forecasting: The Case of Corporate Credit Score Prediction," *Korea Association of Information Systems*, vol. 29, no. 1, 2020, pp. 241-265.

[5] C. Im, "EEG Dimensional Reduction with Stack AutoEncoder for Emotional Recognition using LSTM/RNN," *J. of the Korea Institute of Electronic Communication Science*, vol. 15, no. 4, 2020, pp. 717-724.

[6] C. Neocleous and C. Schizas, "Artificial Neural Network Learning: A Comparative, Methods and Applications of Artificial Intelligence," *Hellenic Conference on Artificial Intelligence SETN, Springer*, Thessaloniki, Greece, 2002, pp. 300-313.

[7] Y. Milad, Y. Moslem, F. Masood, and F. Flavio, "Patient visit forecasting in an emergency department using a deep neural network approach," *The international journal of cybernetics*, vol. 49, issue. 12, 2019, pp. 2335-2348.

[8] M. Kim, "A Study on the Sports Rehabilitation Treatment for the Intellectual Disabilities using deep learning," *J. of the Korea Institute of Electronic Communication Science*, vol. 15, no. 4, 2020, pp. 725-732.

[9] C. Jung, R. Jang, D. Nyang, and K. Lee " A Study of User Behavior Recognition-Based PIN Entry Using Machine Learning Technique," *Korea Information Processing Society review, computer and communication systems*, vol. 7, no. 5, 2018, pp. 127-136.

Author

조준모(Jun-Mo Jo)

1991년 아이오아주립대학교 컴퓨터과학과 졸업 (공학사)
1995년 경북대학교 대학원 컴퓨터공학과 졸업(공학석사)
2004년 경북대학교 대학원 컴퓨터공학과 졸업 (공학박사)
1998년~현재 동명대학교 전자공학과 교수
※ 관심분야 : 인공지능(딥러닝), 챗봇