ISSN 1746-7659, England, UK
Journal of Information and Computing Science
Vol. 16, No. 1, 2021, *pp*.041-051

A Multi-input Time Series Prediction Model Based on CNN-BLSTM

Ting Xiao
Nanjing University of Information Science and Technology, Nanjing-210044, China
(Received January 25, 2021, accepted March 12, 2021)

Abstract. Real time series data sets are often composed of multiple variables. For the future trend of data, not only the historical value of the variables but also other implicit influence factors should be considered. In this paper, a deep neural network prediction model based on multivariable input multi-step output named CNN-BLSTM is proposed. CNN-BLSTM is mainly composed of convolutional neural network (CNN) and bi-directional long short memory network (Bi-LSTM). CNN is used to extract spatial features between variables of multivariate raw data, and Bi-LSTM is used to extract and encode features in time direction. The proposed CNN-BLSTM is able to predict the temperature based on a real-life meteorological data. The experimental results show that the prediction accuracy of the proposed CNN-BLSTM model is significantly better than several state-of-the-art baseline methods.

Keywords: time series, bi-directional long-short term memory, convolutional neural network, prediction, multivariable input

1. Introduction

Time series refers to the ordered set of some variables according to the sequence of time occurrence, which contains the general law that this variable changes with time. Such as monitoring data in meteorological management [1], household electricity consumption data [2], transaction data in the stock market [3], etc. In real life, the collected data is usually multivariate time series (MTS), and there may be complex dynamic interdependencie between different sequences. For example, road surface temperature changes are affected by factors such as temperature, rainfall and humidity. These interdependencies are important, but they are difficult to capture and analyze. The research shows that if the intrinsic correlation and historical characteristics of multivariate time series can be correctly mined, the accuracy of prediction variables can be significantly improved [3].

Researchers have come up with some ways to solve prediction problems. For example, the traditional statistical method integrated automatic regression moving average (ARIMA) model performs well in linear time series data forecasting [4], such as Kumar and Jain [5] used ARIMA to establish a model for predicting traffic noise time series, Ediger and Akar [6] used ARIMA model to predict Turkey's fuel demand and so on. ARIMA is good at predicting stationary time series data. However, for nonstationary time series data, ARIMA faces the problem of increasing smooth step size and difficulty. With the rapid development of machine learning, machine learning methods represented by artificial neural networks (ANN) have been widely used in the classification and prediction of time series because of its strong ability of nonlinear data processing and noise suppression [7,8,9,10]. However, ANN [11] is not as good as ARIMA in the low frequency range of data, and it is easy to produce overfitting in the high frequency range [12].

Recently, deep learning, an advanced machine learning in the field of artificial intelligence, is becoming more and more popular. It combines several machine learning methods to extract high-level expression from complex data by hierarchical learning process and using structure composed of multiple nonlinear transformations [13]. Due to the high computational performance, deep learning models have shown great advantages in automatically extracting and learning multivariate data features in large amounts of data, and have been successfully applied in many scientific fields, including computer vision [14], image classification [15,16], natural language processing [17,18], etc.

Long short-term memory network (LSTM) and convolutional neural networks (CNN)) are two popular deep learning models. LSTM is a recurrent neural network that collects extended sequential data for processing, representation and storage in hidden memory. It solves the problem that RNN is easy to produce gradient explosion or disappearance, which leads to its performance degradation. It becomes a state-of-the-art predictive model to process sequence data in various applications [19], including visual recognition and description [20], continuous Speech recognition [21,22] etc. However, the actual data changes usually depend not only on the previous historical data sequence, but also on the later generated sequence. Elsheikh [23] et al. used a bidirectional LSTM (Bi-LSTM) algorithm to increase the input of LSTM by inputting one step of time series data into the network forward and backward respectively. Compared with LSTM, Bi-LSTM improves the training mechanism to capture information from the past data context while also capturing information from the future data context, so as to obtain more complete information and make predictions with higher accuracy.

CNN [24] is a biologically-inspired type of deep neural network (DNN) that has recently gained popularity due to its success in classification problems (e.g. image recognition [25] or time series classification [26]). The CNN consists of a sequence of convolutional layers, the output of which is connected only to local regions in the input. This is achieved by sliding a filter, or weight matrix, over the input and at each point computing the dot product between the two (i.e. a convolution between the input and filter). This structure allows the model to learn filters that are able to recognize specific patterns in the input data. Studies have shown that the combination of CNN and LSTM can be used to develop more accurate prediction models [27]. For example, Huang et al. [28] found that it is advantageous to use CNN for feature extraction and then input the feature value into the LSTM architecture. In the hybrid CNN-LSTM model, a one-dimensional CNN extracts the deep features of the main elements. Then, use the LSTM model to make predictions using these depth features. Such as Li et al. [29] use a joint CNN-LSTM model to predict collision risk, Barzegar R et al. [30] predict short-term water quality and so on.

Through the above discussion, this paper proposed a multi-variable input-multi-step output time series forecasting model. The model mainly includes two parts: a spatial feature extractor composed of CNN and a temporal feature encoder composed of BLSTM. Before the temporal feature encoder, the spatial feature extractor is used to extract the horizontal relationship of multi-dimensional variables, and then the encoder is used to learn the temporal relationship features obtained by the feature extractor, and make predictions accordingly.

The rest of this article is organized as follows: The second part is mainly an introduction to related neural networks. The third part is a detailed introduction to the framework and algorithm of the proposed CNN-BLSTM model. The fourth part is experiment and result analysis. The fifth part is summary and outlook.

2. Related Networks

2.1 Convolutional Neural Network

CNN is a neural network proposed by LeCun et al. [31] to effectively extract the original image features by simulating the perception mechanism of human natural vision. As shown in Figure 1, each node of the CNN is only connected to a local area in the input. The spatial extent of this connectivity is called the receptive domain of the node. Local connectivity is achieved by replacing the weighted sum of neural networks with convolution. In each layer of the convolutional neural network, the input is convolved with the weight matrix (also called the filter) to create a feature map. Unlike conventional neural networks, all values in the output feature map have the same weight. This means that all nodes in the output detect exactly the same pattern. The local connectivity and shared weights of neural networks reduce the total number of learnable parameters and improve training efficiency. Therefore, the essence behind the convolutional neural network is to learn a weight matrix in each layer so that it can extract the necessary translation invariant features from the input.

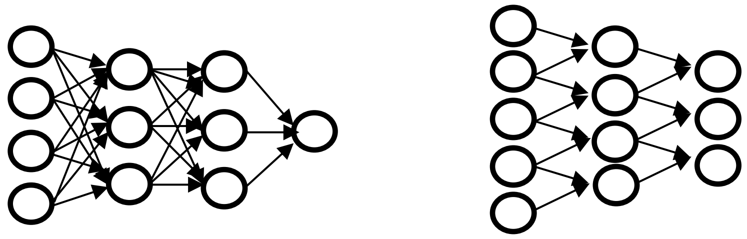

Fig. 1. A feedforward neural network with three layers (L) and a convolutional neural network with two layers and filter size 1 × 2(R).