Predict Music Playback with LSTM Neural Network

Zehua Zhao

School of Software Engineering, Chongqing University, Chongqing 400044, China

1 RNN

The regression problem is very popular and there are many regression models now, such as Linear Regression, Support Vector Regression (SVR), Gradient Boosted Regression (GBRT). But these classic regression models above all have a limitation, since the input of these models must be fixed-length and they don't have sequential structure, these methods have difficulty in capturing the sequential dependency of the variable-length data.

Recurrent Neural Network (RNN) is an effective approach to sequential prediction, which can use their memory information to process sequences of inputs. Now many studies leverage RNN to model the temporal dependency within the data, and most of them get the state-of-the-art results on many practical tasks.

2 LSTM

Long Short Term Memories (LSTM) is a variation of RNN. It replaces the activation function with the LSTM unit, which can process time series with very long time lags of unknown size between important events. It outperforms traditional RNN in numerous applications, such as handwriting recognition and speech recognition.

3 Dataset Description

This dataset has two files : user_actions and songs, the description of these two files is provided in Fig 1 and Fig 2. User_actions file contains 5652231 records in 6 months (20150301-20150830), and songs file contains 26958 songs' information.

We choose the records from 20150301 to 20150630 as training dataset, and the records from 20150701 to 20150831 as testing dataset.

4 Data Preprocessing

We create 50 models to represent the 50 artists respectively. Each artist have 4 lists (ordering in time) to represent the amount of play, collection, download and the initial play of a new song, respectively.

We will use the other 3 lists to predict the play list with LSTM.

列名	类型	说明
user_id	String	用户唯一标识
song_id	String	歌曲唯一标识
gmt_create	String	用户播放时间（unix时间戳表示）精确到小时
action_type	String	行为类型：1，播放；2，下载，3，收藏
Ds	String	记录收集日（分区）

Fig. 1: user_actions

列名	类型	说明
song_id	String	歌曲唯一标识
artist_id	String	歌曲所属的艺人Id
publish_time	String	歌曲发行时间，精确到天
song_init_plays	String	歌曲的初始播放数，表明该歌曲的初始热度
Language	String	数字表示1,2,3...
Gender	String	1,2,3

Fig. 2: songs

5 Metrics

The evaluation metrics we used is defined as Eq. (1):

$$F = \sum_{j \in W} (1 - \sigma_j) \times \phi_j \tag{1}$$

W is the set of artists. σ_j and ϕ_j is defined as Eq. (2) and Eq. (3), respectively.

$$\sigma_j = \sqrt{\frac{1}{N} \sum_{k=1}^{N} (\frac{S_{j,k} - T_{j,k}}{T_{j,k}})^2} \tag{2}$$

$$\phi_j = \sqrt{\sum_{k=1}^{N} T_{j,k}} \tag{3}$$

$S_{j,k}$ and $T_{j,k}$ represent the real value and predicted value of the amount of play for user j in date k, respectively

6 Experimental results

The best F of our experiment is 6210.

The experimental results are shown in Fig. 3. Each graph represents an artist, the red line is the predicted value and the blue line is the real value.

It can be seen from the Fig. 3 that the prediction results can roughly simulate the change trend of the real value. But the error of each point is still very large, We think the reasons are as follow:

- The time regularity of our data is not strong. You can do more work during data preprocessing to avoid this problem, to find more attributes of artists which have strong time regularity as the input of LSTM.
- Training data is too small. Each of the lists we used as input only contains 122 values.

7 Reference

[1]Li X, Wu S, Wang L. Blood Pressure Prediction via Recurrent Models with Contextual Layer[C]// International Conference on World Wide Web. International World Wide Web Conferences Steering Committee, 2017:685-693.

[2]http://blog.csdn.net/mylove0414/article/details/56969181

[3]http://blog.csdn.net/songyunli1111/article/details/78513811

[4]http://blog.csdn.net/a819825294/article/details/53487913

[5]https://zhuanlan.zhihu.com/p/24720659?refer=YJango

[6]https://www.jianshu.com/p/9dc9f41f0b29

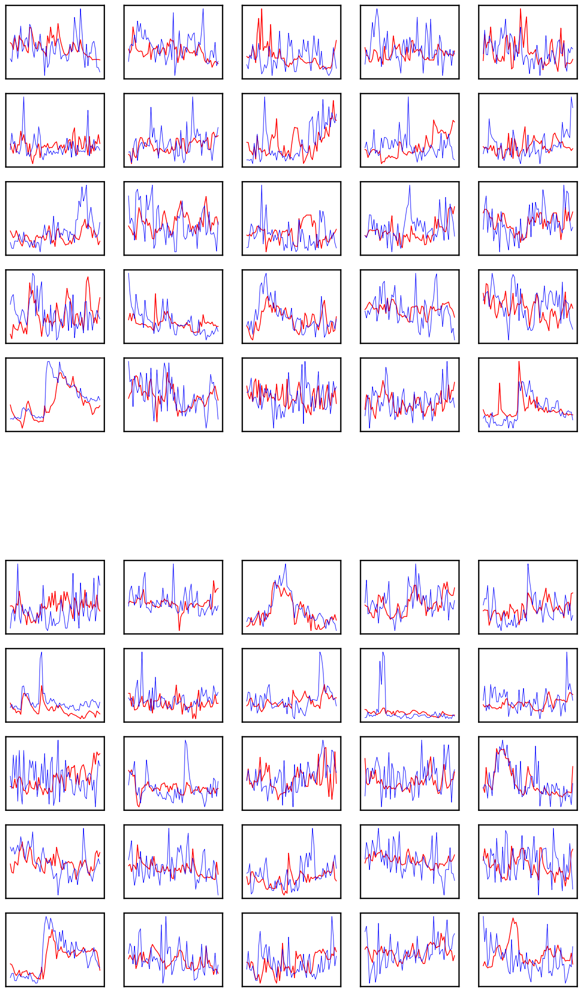

Fig. 3: results