

基于门控深度循环信念网络的边坡沉降预测

武焱, 张映雪

(长沙理工大学 交通运输工程学院, 湖南 长沙 410114)

摘 要:本研究针对现有边坡沉降预测模型精度低、无法有效反映沉降值蕴含的时序信息等问题,提出基于门控深度循环信念网络(GDRBN)的边坡沉降混合预测模型。为提高训练效率,引入自适应学习率,并以广佛肇高速公路二期工程为实例,建立多种边坡沉降预测模型,并进行计算比较。研究表明:基于GDRBN的边坡预测模型的预测精度比GM、BP、RNN、DBN预测模型的分别提高了69%、54%、38%、26%,可为边坡预测提供更准确的计算方法。

关键词:边坡;沉降预测;深度学习;循环神经网络;自适应学习率

中图分类号: TU146.1

文献标志码: A

Slope settlement prediction based on gated deep recurrent belief network

WU Yan, ZHANG Yingxue

(School of Traffic and Transportation Engineering, Changsha University of Science & Technology, Changsha 410114, China)

Abstract: The existing slope settlement prediction models have low accuracy and cannot effectively reflect the time series information contained in the settlement. A hybrid prediction model for slope settlement was proposed based on the gated deep recurrent belief network (GDRBN). The adaptive learning rate is introduced to improve the efficiency of model training. Taking the project of the second phase of Guangfo-Zhaoqing Expressway as an example, different settlement prediction models was established, and the outcomes were contrasted. The results show that the proposed models will increase the accuracy of the prediction by 69%, 54%, 38%, and 26%, respectively, which is compared with traditional prediction models (such as GM, BP, RNN, DBN). It can provide a accurate calculation method for slope prediction.

Key words: side slope; settlement prediction; deep learning; recurrent belief network; adaptive learning rate

公路修筑过程中,因山体开挖或填筑形成了许多边坡,打破了原本土体的力学平衡,且在内部因素(地质缺陷、压缩变形等)和外部因素(降雨等)作用下,极易诱发滑坡、坍塌等灾害^[1]。对于部分高边坡,若不能及时获取其形变情况,将无法做出相应防护措施,进而引发交通事故等二次灾害,造成人员伤亡和财产损失。

在边坡沉降预测中,灰色预测模型(gray forecast model, GM)是通过数学建模获取沉降预测值的经典方法^[2]。但传统GM模型在初值选取和背景值构造过程中产生的误差会降低其预测精

度,为此,有学者提出了改进初值和背景值的GM预测模型^[3]。为精准预测不同场景下的边坡沉降值,在传统GM基础上还衍生出了多变量灰色模型(multi-variable gray model, MGM)、优化灰色模型(optimizing gray model, OGM)、动平均改进GM模型等^[3-5]。但受GM方法的构造约束,其预测精度依然偏低。

随着机器学习和智能算法的发展,边坡沉降预测精度得到极大提升。彭立顺等^[6]利用误差反向传播神经网络(back propagation neural network, BPNN)对大量现场实测数据进行学习,建立边坡沉

降预测模型,极大提高了预测精度,但该模型求解耗时较长。也有研究采用相空间重构理论对边坡位移原始数据进行转化,并将其作为支持向量机(support vector machines, SVM)的输入数据进行预测,建立改进的SVM预测模型,有效降低模型计算时长^[7]。

近年来,许多学者将深度学习方法应用到边坡沉降预测中。循环神经网络(recurrent neural network, RNN)是基于传统神经网络发展而来的,基于隐含层的全连接型结构具有较强的序列处理能力,可有效反映沉降量中蕴含的时序信息^[8]。但RNN在模型生成过程中容易出现梯度消失。长短期记忆神经网络(long short-term memory, LSTM)^[9]和门控循环单元神经网络(gated recurrent unit, GRU)^[10]均为RNN的变形结构,可有效解决沉降预测模型中梯度消失的问题。RNN的应用提高了边坡沉降预测精度,但其结构仍是浅层网络,而深度信念网络(deep belief network, DBN)通过叠加多个受限玻尔兹曼机(restricted boltzmann machine, RBM),经过多层结构间数据的传递学习和误差调节,可极大提高沉降预测精度^[11]。但DBN弱化了沉降序列中的时序性。

门控深度循环信念网络(gated deep recurrent belief network, GDRBN)利用GRU替代RBM中的隐含层信息,有效融合了GRU的时序处理能力和DBN的数据挖掘能力^[12]。因此,本研究拟将GDRBN应用于边坡沉降预测,并基于自适应学习率的模型提出改进算法。以广佛肇高速公路二期工程为实例,建立边坡沉降预测模型,并将所得结果与GM、BP、RNN、DBN模型预测结果进行对比,验证本模型的预测精度及时序处理的有效性。

1 循环神经网络与深度信念网络基本模型

1.1 循环神经网络基本模型

针对传统反向传播神经网络(back propagation neural network, BPNN)在处理数据时无法有效反映输入数据隐含的时间信息问题,ELMAN等学者于20世纪80年代末提出了RNN。RNN在层次结构上与BPNN相同,均由输入层、隐含层和输出层组成。其主要区别为:BPNN隐藏层的神经元间无连接,而RNN隐藏层内的各神经元间为全连接。这种特殊结构,使隐藏层在任意时刻输出的数据不仅与当前时刻的输入数据有关,还与之前时刻隐藏层的历史输出数据有关。研究表明:RNN对输入数据具有记

忆功能,可有效反应数据当中的时序信息。RNN模型隐藏层标准结构如图1所示。

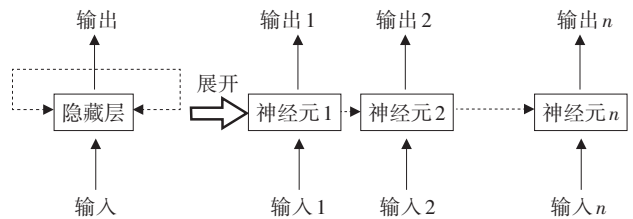


图1 RNN模型隐藏层标准结构

Fig. 1 Standard structure of hidden layer of the RNN model

RNN结构增加了时序数据的处理能力,但当隐藏层神经元数量过多时,位于时间轴前段的数据无法准确传递到时间轴后端的神元。在实际应用中,RNN的记忆功能只能维持若干个时段,即会出现梯度消失的现象。SEPP等学者于1997年提出了基于门控制的LSTM。LSTM在隐藏层中增加了遗忘门、输入门和输出门3个门控制函数,通过自动修改权值系数,控制不同门的激活,实现数据的长期记忆,有效避免RNN模型中梯度消失情况的出现。

LSTM中的遗忘门负责筛选上一神经元输出数据,遗忘次要信息,仅将重要的数据保留并传递给输入门;输入门负责当前输入数据的更新及经遗忘门筛选过的数据的输入,起到信息传递的作用;输出门负责神经元的输出及隐藏层时序信息的返回。

LSTM通过3个门的巧妙配合增加了RNN的记忆能力,但模型更新逻辑复杂,计算量大。CHO等学者于2014年提出了LSTM的简化模型——GRU。GRU将LSTM的遗忘门和输入门合并为更新门,负责对历史信息的筛选和新信息的更新,引入复位门代替输出门,负责控制历史信息对当前状态的贡献,模型如图2所示。LSTM、GRU门的激活仅由当前的输入与先前输出共同决定,GRU门的激活则无需额外存储单元,更新逻辑更简单,计算量更小。

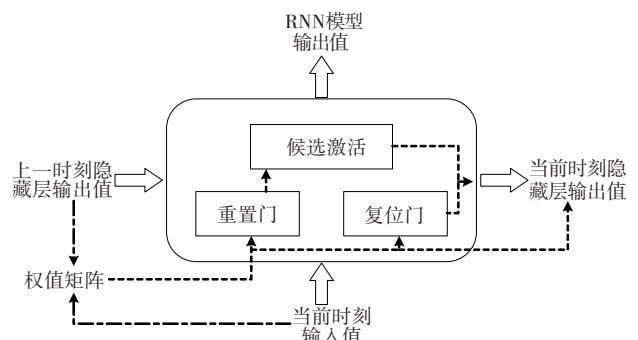


图2 GRU隐藏层神经元模型

Fig. 2 Neuron model of hidden layer of the GRU

1.2 深度信念网络基本模型

DBN是一种基于概率的典型深度学习网络模型,相比浅层网络其具有更强的数据挖掘和预测能力。在结构上,DBN由数个仅含可视层和隐含层的RBM堆叠而成,下层RBM隐藏层和上层RBM可视层共用相同神经元,并经单层BP神经网络输出。

RBM是一种基于概率的两层无向图模型,其可视层 v 和隐含层 h 中的每个神经元均服从0-1分布,且同层各神经元间无任何连接,层间各神经元通过权重矩阵实现全连接。RBM的基本结构如图3所示。RBM单个神经元的激活概率可参考文献[11]。

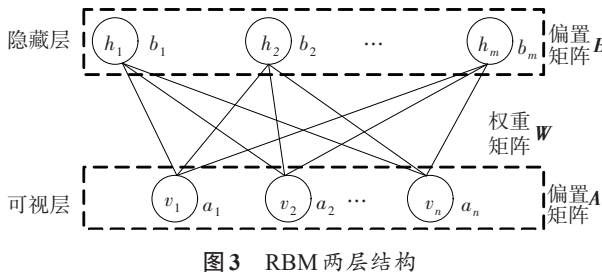


图3 RBM两层结构

Fig. 3 Two-layer structures of the RBM

含3个RBM单元的DBN基本结构如图4所示。DBN先通过底层RBM1的可视层读取数据;再采用对比散度算法(contrastive divergence algorithm, CD)由下到上对RBM进行逐层无监督训练,求解各层权重矩阵和偏置系数初值;然后采用梯度下降算法或反向传播算法由上向下对训练过程得到的权重矩阵和偏置系数进行微调,得到最优网络参数;最后,通过顶层BP神经网络输出预测结果。

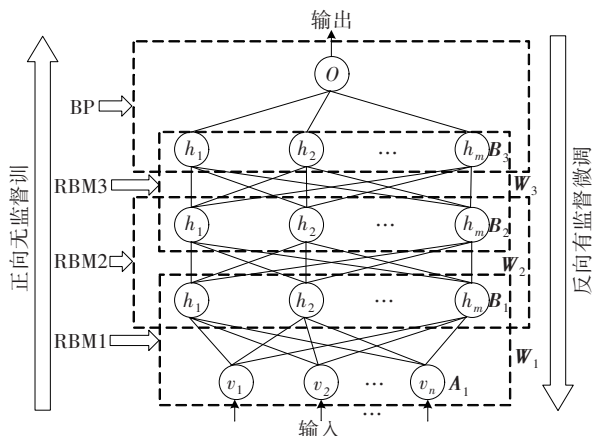


图4 DBN结构

Fig. 4 Structure diagram of the DBN

2 基于GRU和DBN的混合预测模型

2.1 门控深度循环信念网络基本模型

传统DBN模型虽具有较强的数据挖掘能力,但不能很好地提取数据信息中隐含的时序关系。GRU通过重置门、更新门和隐藏层神经元全连接的特殊结构,实现了对历史数据的记忆功能,能较好地反映信息中的时序关系,但其浅层结构限制了预测精度的提升。因此,提出了一种兼顾GRU时序处理能力和DBN数据预测能力的混合预测网络模型——GDRBN^[12]。

先将GRU蕴含的时序信息替换RBM的隐含层信息,组成GRU-RBM单元,然后,堆叠数个GRU-RBM单元形成GDRBN。假设 K 层GDRBN模型的输入信息为 $\{v_1, \dots, v_t, \dots, v_T\}$ (T 为输入数量), k 层GRU在 t 时段的隐藏层状态为 $g_{k,t}$, k 层RBM在 t 时段的可视层和隐藏层状态为 $v_{k,t}$ 和 $h_{k,t}$,且 $v_{k,t}, h_{k-1,t}$ 为同一神经元在相邻两层GRU-RBM的不同表示方式,模型输出层为 Y_t ,最终输出预测数据为 \hat{Y}_t ;记第 k 层中相邻时刻GRU的连接权重矩阵为 W_k^{gg} ,第 k 层中GRU与其隐藏层神经元间的连接权重矩阵为 W_k^{gh} ,第 k 层中GRU与其可视层神经元间的连接权重矩阵为 W_k^{gv} ,第 k 层中隐藏层神经元与可视层神经元间的连接权重矩阵为 W_k^{hv} ,第 k 层中可视层神经元与GRU间的反馈权重矩阵为 W_k^{vg} ,RBM可视层和隐藏层的偏置矩阵为 A_k, B_k ,最后一层隐含层与输出层的连接权重为 O_t ,则 k 层GDRBN的基本结构如图5所示。

GDRBN在训练时包括正向无监督学习、纵向误差反馈微调 and 横向误差反馈微调3个过程。正向无监督学习通过底层GRU-RBM可视层输入信息,并由下向上逐层优化权重矩阵与偏置矩阵,如图5中白色箭头所示。纵向误差反馈微调在各层GRU-RBM内进行误差传递,以保证GDRBN向最优方向修正参数,如图5中阴影箭头所示。横向误差反馈微调在时间轴上进行误差传递,以保证GDRBN有效保留数据中的时序信息,如图5中灰色箭头所示。通过这3个过程的协调,使GDRBN在保留数据中时序信息的基础上提高了预测精度。

2.2 正向无监督学习过程

正向无监督学习的目的是生成各层GRU-RBM权重矩阵和偏置矩阵,通常采用CD算法。但该算

法往往需要多次迭代且每次参数更新方向均有所不同,若训练过程中保持学习率为定值,可能导致算法不收敛或效率低下,所以制定合适的学习率是无监督学习过程的关键。本研究基于相邻参数更新过程中迭代方向,设计自适应学习率方法。考虑自适应学习率,则 t 时刻的可视层与隐藏层间的权重参数 $\mathbf{W}_{k,t}^{\text{hv}}$ 更新方式为:

$$\mathbf{W}_{k,t}^{\text{hv}} = \mathbf{W}_{k,t-1}^{\text{hv}} + \eta_{k,t} \frac{\partial \ln P(v)}{\partial \mathbf{W}_{k,t}^{\text{hv}}} \quad (1)$$

$$\eta_{k,t} = \begin{cases} U\eta_{k,t-1}, \Delta_{k,t} \times \Delta_{k,t-1} > 0 \\ D\eta_{k,t-1}, \Delta_{k,t} \times \Delta_{k,t-1} < 0 \\ \eta_{k,t-1}, \Delta_{k,t} \times \Delta_{k,t-1} = 0 \end{cases} \quad (2)$$

$$\Delta_{k,t} = v_{k,t}h_{k,t} - v_{k,t-1}h_{k,t-1} \quad (3)$$

式中: $P(v)$ 为可视层神经元被激活的概率函数,可由文献[11]获得; $\eta_{k,t}$ 为 t 时刻的学习率; U 、 D 分别为自适应学习率增加、减小系数,且满足 $0 < D < 1 < U$; $\Delta_{k,t}$ 为 t 时刻参数的更新方向,由 t 和 $t-1$ 时刻可视层与隐藏层的神经元概率共同决定。

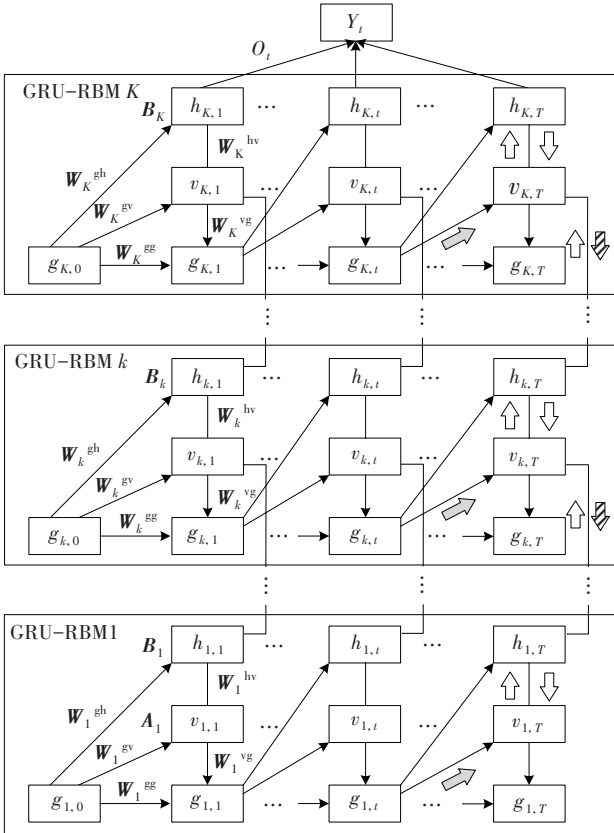


图5 GDRBN结构与信息传递过程

Fig. 5 Structure and information transferring process of GDRBN

t 时刻GRU-RBM单元内偏置矩阵参数 $\mathbf{A}_{k,t}$ 、 $\mathbf{B}_{k,t}$ 和GRU隐藏层输出 $g_{k,t}$ 不仅与初始输入数据

相关,还与上一时刻历史数据相关,其参数更新方式为:

$$\begin{aligned} \mathbf{A}_{k,t} &= \mathbf{A}_{k,0} + \mathbf{W}_{k,t}^{\text{gv}} g_{k,t-1} \\ \mathbf{B}_{k,t} &= \mathbf{B}_{k,0} + \mathbf{W}_{k,t}^{\text{gh}} g_{k,t-1} \end{aligned} \quad (4)$$

$$g_{k,t} = \sigma[g_{k,t} + \mathbf{W}_{k,t}^{\text{gg}} g_{k,t-1} + \mathbf{W}_{k,t}^{\text{vg}} g_{k,t-1}]$$

式中: $\mathbf{A}_{k,0}$ 、 $\mathbf{B}_{k,0}$ 、 $g_{k,0}$ 分别为随机初始参数; $\mathbf{W}_{k,t}^{\text{gv}}$ 、 $\mathbf{W}_{k,t}^{\text{gh}}$ 分别为 t 时段第 k 层GRU与其可视层神经元、GRU与其隐藏层神经元的连接权重矩阵, $\mathbf{W}_{k,t}^{\text{gg}}$ 为 t 时段第 k 层GRU间的连接权重矩阵, $\mathbf{W}_{k,t}^{\text{vg}}$ 为 t 时段第 k 层可视层神经元与GRU间的反馈权重矩阵,由式(1)计算得到; $\sigma[\cdot]$ 为sigmoid函数。

在获得各层权重矩阵与偏置矩阵的参数更新方式后,GDRBN的正向无监督学习过程可写为:

$$S_{k,t,i} = \begin{cases} \sigma[\mathbf{W}_{k,t,i}^{\text{vg}} v_{k,t} + b_{k,t,i}], k=1 \\ \sigma[\mathbf{W}_{k,t,i}^{\text{vg}} S_{k,t,i-1} + b_{k,t,i}], k \in [2, \dots, K-1] \\ \sigma[\mathbf{W}_{k,t,i}^{\text{vg}} S_{k,t,i-1} + \mathbf{W}_{k,t,i}^{\text{gg}} S_{k,t,i} + b_{k,t,i}], k=K \end{cases} \quad (5)$$

$$\hat{Y}(t) = f[\mathbf{O}_t S_{K,t,i} + b_{K,t,i}] \quad (6)$$

式中: $S_{k,t,i}$ 为 t 时段第 k 层第 i 个隐藏层神经元输出值; K 为隐含层的总层次; $\mathbf{W}_{k,t,i}^{\text{vg}}$ 、 $\mathbf{W}_{k,t,i}^{\text{gg}}$ 分别为权重矩阵 $\mathbf{W}_{k,t}$ 、 $\mathbf{W}_{k,t}$ 中第 i 个列向量; $b_{k,t,i}$ 为偏置矩阵 $\mathbf{B}_{k,t}$ 中第 i 个列向量; $f[\cdot]$ 为输出层激活函数。

2.3 纵横误差反馈微调过程

GDRBN模型 t 时段的预测误差 E_t 为模型预测值 \hat{Y}_t 与模型输出值 Y_t 之差,计算式为:

$$E_t = \hat{Y}_t - Y_t \quad (7)$$

由式(8)可知,预测误差仅和RBM可视层与GRU隐含层间的权重系数 $\mathbf{W}_{k,t,i}^{\text{vg}}$ 有关,引入中间变量函数 $\psi_{k,t,i}$ 描述隐藏层输出与权重 $\mathbf{W}_{k,t,i}^{\text{vg}}$ 的关系,其表达式为:

$$\psi_{k,t,i} = \mathbf{W}_{k,t,i}^{\text{vg}} S_{k,t,i-1} + b_{k,t,i} \quad (8)$$

GDRBN纵向误差为预测误差关于中间变量的一阶偏导数,其值在各层GRU-RBM之间反馈传导。由链式求导法可得纵向误差 $\bar{E}_{k,(i-1)}$:

$$\bar{E}_{k,(i-1)} = \frac{\partial E_t}{\partial \psi_{k,t,i-1}} = \frac{\partial E_t}{\partial \psi_{k,t,i-1}} \prod_{m=1}^i \frac{\partial \psi_{k,m,(i-1)}}{\partial \psi_{k,m,(i-1)}} \quad (9)$$

GDRBN横向误差为预测误差关于中间变量的二阶偏导数,其值延时间轴反馈传导。由链式求导法可得横向误差 $\underline{E}_{k,(i-1)}$:

$$\underline{E}_{k,(i-1)} = \frac{\partial^2 E_t}{\partial \psi_{k,t-1,i-1}^2} = \frac{\partial E_t}{\partial \psi_{k,t-1,i-1}} \frac{\partial \psi_{k,t-1,i-1}}{\partial \psi_{k,t-1,i-1}} \quad (10)$$

本研究采用均方差(mean squared error, MSE)作为统计模型输出误差状况的损失函数,以避免模

型在反馈微调过程中陷入局部最优。

2.4 基于GDRBN的边坡沉降预测模型

边坡沉降值一般随着时间推移逐渐递增,其历史检测沉降信息中包含大量时序信息,而GM、BP神经网络等预测模型均无法有效反映时序与沉降值间的隐含关系。因此,本研究建立基于GDRBN的边坡沉降预测模型。

针对工程现场检测条件及存储不当导致的数据时距不等与缺失的问题,采用三次样条插值法进行数据插补。为加快GDRBN网络收敛速度,避免梯度爆炸,采用min-max标准化方法将插补后的数据进行归一化,完成原始数据预处理。将预处理后的数据分为训练与检测数据两部分,其中,训练数据用于确定GDRBN网络结构与参数,检测数据用于验证模型有效性。将训练数据集输入GDRBN模型,反复优化GDRBN模型参数直至达到设定精度,输出最优权重矩阵与偏置矩阵,获得优化的GDRBN模型。再将检测数据集输入优化的GDRBN模型,并输出预测结果。通过计算模型预测结果与实际沉降值间的平均绝对误差百分比 e_{MAPE} 、相关系数 C 和决定系数 R^2 的3个指标^[10],对GDRBN模型预测结果进行分析。GDRBN模型的边坡沉降预测流程如图6所示。

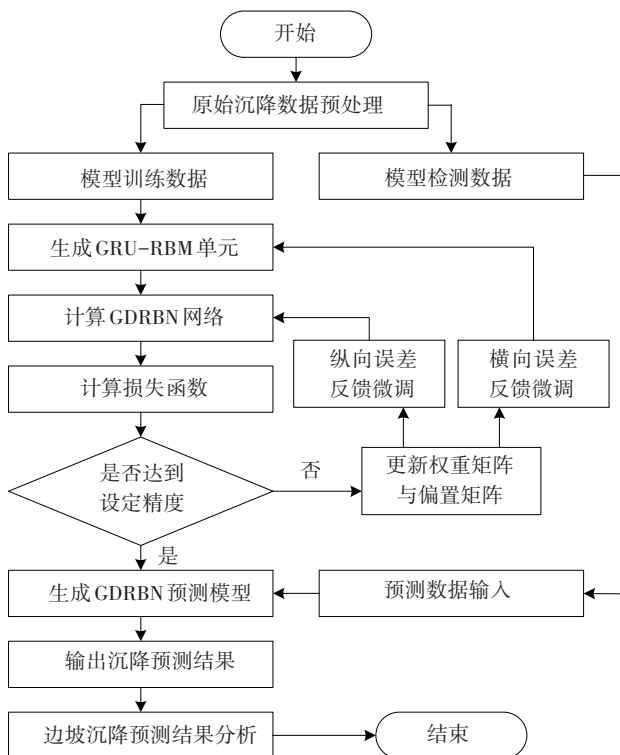


图6 基于GDRBN模型的边坡沉降预测流程

Fig. 6 Slope settlement prediction process based on GDRBN model

3 工程实例

3.1 基础数据

以广佛肇高速公路二期肇庆大旺至封开江口段K171+005~K171+185路段工程为实例,建立模型进行计算。该段线路全长为180 m,属于高路堑边坡,沿线多为风化石、高液限黏土,土质状况较差,采用全站仪监测各位移监测点的坐标和高程,通过对比测量数据监测边坡变形情况。根据该段坡高、坡长及岩土体情况,从挖方边坡最高处开设监测断面,并沿路线方向每隔30~50 m间距向两侧均匀布置地表变形监测断面,共布置3个监测断面,6个监测点。变形监测点平面、横断面如图7~8所示,图中A、B、C表示3个监测点。

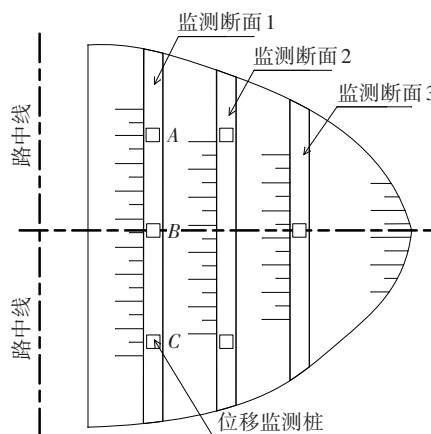


图7 变形监测点平面示意

Fig. 7 Plane diagram of the deformation monitoring point

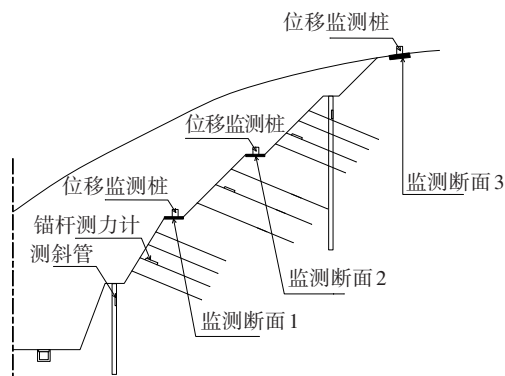


图8 变形监测点横断面示意

Fig. 8 Cross section diagram of the deformation monitoring point

取监测断面1居中的监测点B在99~360 d的20个实测数据为原始沉降数据,见表1。使用Matlab三次样条插值语句spline对原始沉降数据序列进行数据

填补与等时距处理,以 9 d 为一监测周期,共计获得 30 个周期的沉降数据序列。其中,前 25 个周期数据用于 GDRBN 模型训练,后 5 个周期数据用于模型预测。经 min-max 标准化后的数据见表 2。插值后的最大沉降值为 23.653 mm,最小沉降值为 18.851 mm。

表 1 原始沉降序列

Table 1 Original settlement sequence

天数/ d	沉降 值/mm	天数/ d	沉降 值/mm	天数/ d	沉降 值/mm	天数/ d	沉降 值/mm
99	21.8	131	23.3	165	19.3	270	21.3
105	22.4	139	17.4	171	20.7	284	22.6
111	23.6	147	20.3	178	21.3	309	20.8
117	19.2	151	20.5	208	22.1	333	19.4
125	21.3	158	20.3	239	21.8	360	22.4

表 2 预处理后的边坡沉降量

Table 2 Slope settlement after pretreatment

序号	沉降值/mm	序号	沉降值/mm	序号	沉降值/mm
1	0.614	11	0.567	21	0.702
2	1.000	12	0.623	22	0.794
3	0.073	13	0.672	23	0.684
4	0.644	14	0.698	24	0.477
5	0.304	15	0.694	25	0.281
6	0.000	16	0.654	26	0.153
7	0.339	17	0.574	27	0.114
8	0.143	18	0.482	28	0.185
9	0.385	19	0.439	29	0.386
10	0.518	20	0.510	30	0.739

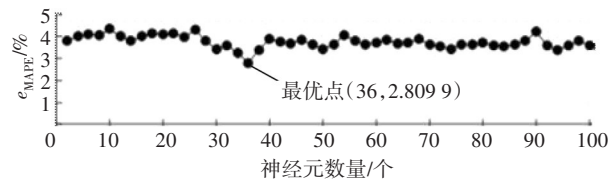
3.2 GDRBN 模型网络结构确定

GDRBN 预测模型的超参数主要包括输入层神经元个数、隐含层层数(即 GRU-RBM 单元个数)、隐含层神经元个数和学习率。

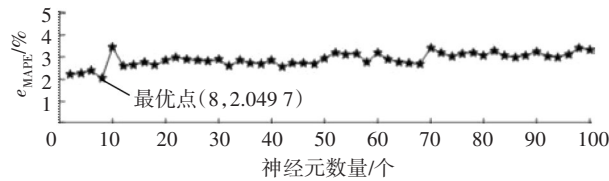
影响边坡沉降的因素主要有地基处理方式、软土层厚度、软土层压缩模量、路堤高度、路基填筑时间和时间序列编号,共 6 个输入变量。因此,输入层神经元个数可设定为 6。GRU-RBM 单元学习率设定为 0.01,自适应学习率增加、减小系数分别设定为 1.4、0.7,迭代次数设定为 300。

GDRBN 隐含层层数及神经元个数通过试验法确定。在给定训练数据集、输入神经元个数、学习率和迭代次数的条件下,分别训练 1~5 层隐含层和 1~100 个神经元节点的排列组合预测模型,并记录下每组模型的 e_{MAPE} 值,根据最优值确定隐含层层数与神经元个数。不同 GDRBN 结构下

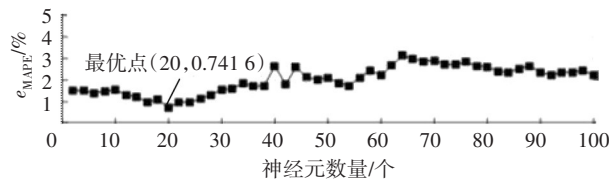
的训练误差如图 9 所示。



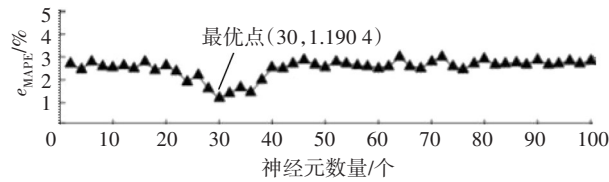
(a) 1 层隐含层 GDRBN 训练结果



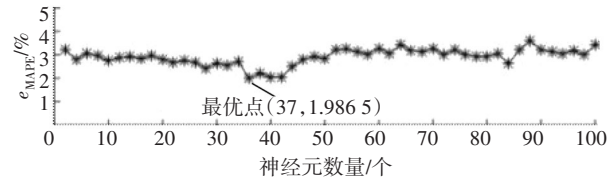
(b) 2 层隐含层 GDRBN 训练结果



(c) 3 层隐含层 GDRBN 训练结果



(d) 4 层隐含层 GDRBN 训练结果



(e) 5 层隐含层 GDRBN 训练结果

图 9 不同 GDRBN 结构下的训练结果

Fig. 9 Training results under different the GDRBN structures

从图 9 可以看出,在不同层数 GDRBN 模型训练过程中, e_{MAPE} 值基本在其均值上下波动,且随着隐含层层数的增加, e_{MAPE} 均值先减小后增大;在层数确定情形下,随着神经元数量的增加,模型 e_{MAPE} 值整体上呈现先减小后增大的变化趋势,局部呈现小范围波动的趋势。这是由于隐含层层数和层内神经元数量初步增加提高了模型的数据挖掘能力,进而提高了预测精度。但随着层数和神经元数量的进一步增加,模型计算的复杂度和误差积累越来越大,导致模型预测性能下降。当隐含层层数为 3、层内神经元数量为 20 个时,模型训练误差最小、性能最佳。因此,本算例 GDRBN 预测模型的最优网络结构确定为 6-20-20-20-1。

3.3 模型预测结果及分析

将 GDRBN 模型应用于广佛肇高速公路二期肇

庆大旺至封开江口段K171+005~K171+185路段的边坡沉降预测,并将计算结果与传统预测模型GM、浅层网络预测模型BP、RNN及DBN的进行对比。通过计算各模型的 e_{MAPE} 、 C 、 R^2 指标,评价各模型的性能。其中,GM模型的发展系数和灰作用量分别为-0.018 5、20.793,BP、RNN、DBN网络结构采用试验法确定,其结构分别为6-25-1、6-21-1、6-20-20-1。学习率设定为0.01,迭代次数设定为300。

GDRBN与其他模型预测结果对比如图10所示。各模型预测值与实测值对比如图10(e)所示(基准线为 $y=x$ 的一次函数,表示预测值与实际值完全相等,图中数据点越靠近基准线,表示其预测精度越高,预测误差越小),各模型预测结果相对误差对比如图10(f)所示,各模型预测结果评价指标见表3。

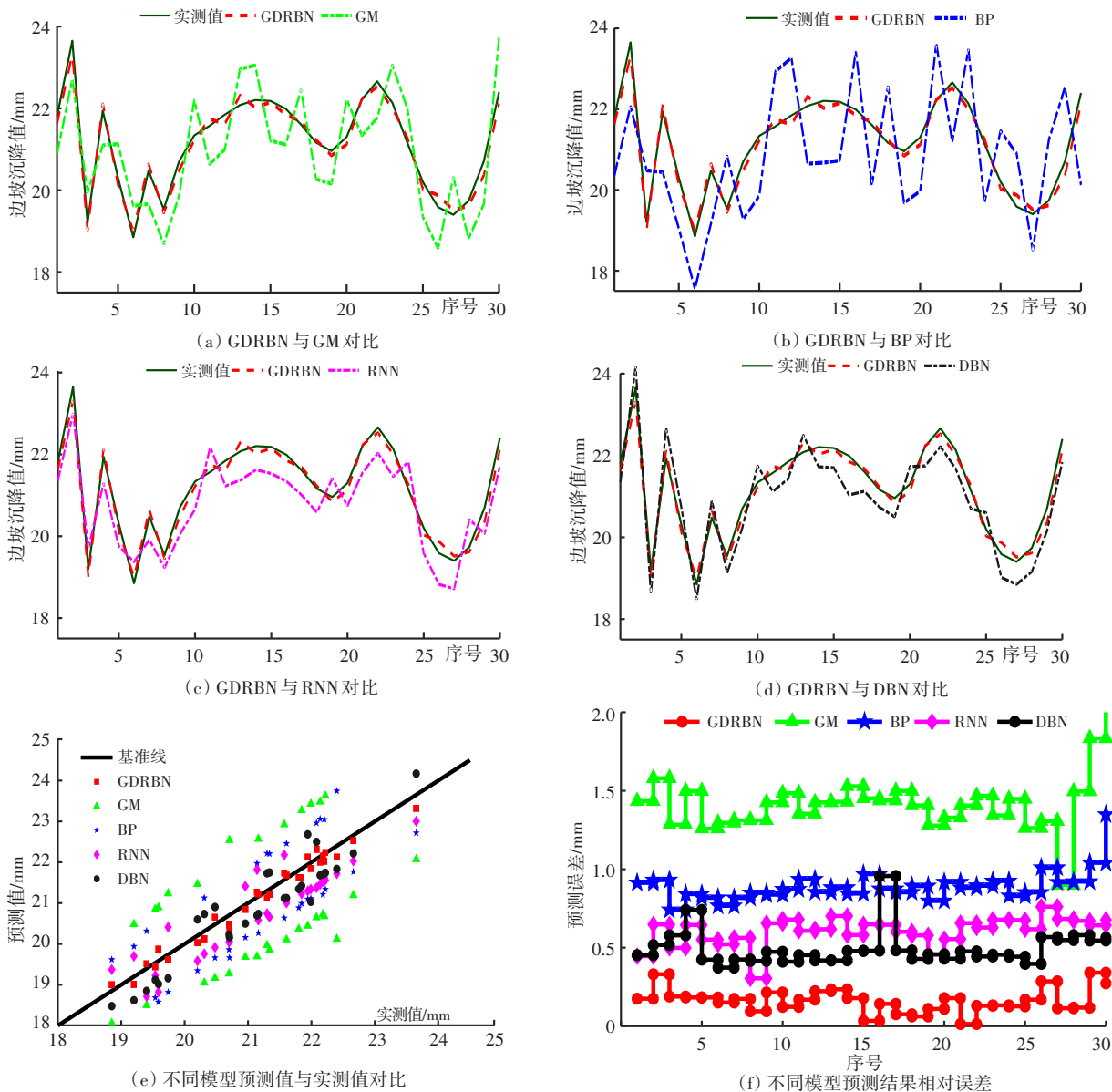


图10 不同模型预测结果

Fig. 10 Prediction results of different models

表3 不同模型的预测结果评价指标

Table 3 Evaluation indicators for prediction results of different models

	$e_{MAPE}/\%$	相关系数 C	决定系数 R^2
GDRBN	2.34	0.935	0.918
GM	7.58	0.477	0.228
BP	5.12	0.675	0.400
RNN	3.75	0.887	0.746
DBN	3.16	0.814	0.853

从图 10 和表 3 可以发现,GM 模型预测结果最差,BP 网络的其次,RNN 与 DBN 的接近,GDRBN 的各项指标最优。GM 模型通过对原始数据建立微分方程达到预测的效果,适用于呈指数增长的数据的预测,而边坡沉降值随着时间的推移有增有减,指数增长趋势不明显。因此,其预测效果最差,各项指标均不如机器学习预测模型的。BP 预测模型相较于 GM 模型,无须建立精确数学模型,对数据变化趋势亦无特殊要求,BP 预测模型通过挖掘数据内部的潜在规律进行预测,相比 GM 预测误差降低了 32.45%。RNN 和 DBN 模型的 e_{MAPE} 值近似,但在相关系数 C 和决定系数 R^2 上各有优劣。相关系数 C 用于评价预测序列在时间维度上的准确性,其值越大,预测数据对时间序列的反应程度越高;决定系数 R^2 用于评价预测模型结果方面的准确性,其值越大,误差越小,模型数据挖掘能力越强,距离回归基准线越近。由表 3 可知,RNN 具有较大的 C 和较小的 R^2 ,而 DBN 具有较小的 R^2 和较大的 C 。从图 10 (e) 还可以看出,DBN 更接近基准线,这表明:RNN 具有较强的时序处理能力,DBN 具有较强的数据挖掘能力。本研究提出的 GDRBN 模型,兼顾了 RNN 的时序处理能力和 DBN 的数据挖掘能力,极大提高了模型预测精度,其 e_{MAPE} 低至 2.34%,比 GM 的和 DBN 的分别降低了 69.13% 和 25.95%,其相关系数 C 为 0.935,相比 RNN 和 DBN 的分别提高了 5.41% 和 14.86%;决定系数 R^2 为 0.918,相比 RNN 和 DBN 的分别提高了 23.06% 和 7.62%。因此,GDRBN 预测模型通过对原始沉降序列中的时序信息进行深度挖掘,有效提高了沉降预测精度,在边坡沉降预测方面具有较强的适用性。

3.4 自适应学习率与固定学习率对比

为验证本模型自适应学习率在模型训练过程中的有效性,将固定低学习率和固定高学习率的 GDRBN 模型结果与自适应学习率 GDRBN 模型预测结果进行对比,数据见表 4,对比如图 11 所示。其中,低学习率取为 0.001,高学习率取为 0.1,自适应学习率的学习效率设定为 0.01,自适应学习率增加、减少系数分别设定为 1.4 和 0.7,

迭代次数设定为 300。

从图 11 可以看出,低学习率的 GDRBN 模型损失函数值,在模型训练约为 320 次时可收敛至 0.122,且随着训练次数的增加,损失函数值呈小幅度下降的趋势,其训练效率较低,耗时较长。当学习率进一步减小时,预测模型可能陷入局部最优。而高学习率的 GDRBN 模型,可以提高训练效率,使模型快速收敛,但在相邻训练过程中调节幅度过大会导致模型无法找到最优值,而是在最优值附近震荡,增加了模型损失函数值。当学习率进一步增大时,预测模型可能会发散。本模型采用自适应学习率,通过动态调节相邻训练过程间的学习速度,可快速提高模型训练效率,找到全局最优点,降低损失函数值。自适应学习率模型损失函数值比低学习率和高学习率模型的分别降低了 42.62%、58.82%,其训练次数和训练耗时适中。

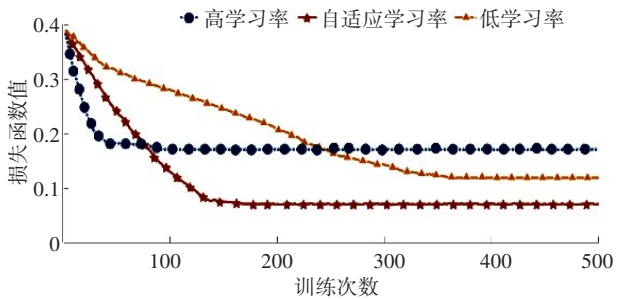


图 11 模型损失随训练次数变化趋势

Fig. 11 The trend of model loss with the number of training

表 4 不同学习率时的模型性能

Table 4 Model performance at different learning rates

不同学习率	$e_{MAPE}/\%$	稳定损失函数值	最小训练次数	训练时间/s
低学习率	1.58	0.122	326	25.21
高学习率	1.97	0.170	88	6.77
自适应学习率	1.11	0.070	175	13.51

3.5 纵横误差反馈微调分析

为验证增加纵横误差反馈微调过程对深度学习网络模型对边坡沉降预测精度的影响,设置仅考虑纵向和横向误差反馈微调过程的 GDRBN 与工作对比。不同处理方法的预测结果评价指标见表 5。

表5 考虑不同误差时的模型预测结果评价指标

Table 5 Evaluation index of model prediction result considering different errors

不同反馈微调	$e_{\text{MAPE}}/\%$	相关系数 C	决定系数 R^2
考虑纵横误差	1.11	0.935	0.928
仅考虑纵向误差	1.35	0.811	0.844
仅考虑横向误差	1.41	0.912	0.754

由表5可知,增加纵横误差反馈微调后的GDRBN模型具有更高的精度、更大的相关系数和决定系数。仅考虑纵向误差反馈微调时,预测结果的相关系数较低,时序处理能力较弱, e_{MAPE} 增加21.62%;仅考虑横向误差反馈微调时,预测结果的决定系数较低,数据挖掘能力较弱, e_{MAPE} 增加27.03%。

4 结论

针对现有边坡沉降预测模型预测精度低、无法反映沉降值时序信息问题,建立了具有时序处理能力的深度学习网络边坡沉降预测模型,得出结论:

1) 通过将GRU蕴含的时序信息替换RBM隐含层信息,组成GRU-RBM单元,叠加数个GRU-RBM单元可形成GDRBN模型。在传统DBN模型正向过程中,引入自适应学习率,加快网络训练效率。在反向微调过程中,引入纵横误差反馈微调网络,有效减少训练误差。GDRBN模型有效融合了RNN模型的时序处理能力和DBN模型的数据挖掘能力,具有较高的预测精度。

2) 基于实际工程样本,利用GDRBN模型对最终沉降值进行预测,模型预测结果较优,沉降值预测值与实测值的平均相对误差为1.11%,相关系数为0.935,决定系数为0.928,相比GM、BP、RNN、DBN等预测模型,其精度更高,误差更小,时序反应能力更优。

3) 受工程实例限制,本模型仅包含30个样本数据,样本过少可能会导致模型无法获得数据中隐藏的真实规律,使得模型在较短时间内预测精度高,而在长时间范围内预测精度大幅下降。因此,可考虑寻找更多样本数据,验证本模型的有效性。本研究所建模型仅考虑边坡沉降序列中

的时序信息,未考虑到相邻监测点沉降量对研究监测点的空间影响。后续可基于目前研究工作,开展考虑沉降量时空相关性的边坡沉降值预测模型。

参考文献(References):

- [1] 刘新喜,邓子君,李盛南,等.炭质泥岩路堑边坡湿化变形模拟分析[J].交通科学与工程,2021,37(1):1-6.
LIU Xinxi, DENG Zijun, LI Shengnan, et al. Simulation analysis of wetting deformation of carbonaceous mudstone cutting slope [J]. Journal of Transport Science and Engineering, 2021, 37(1): 1-6.
- [2] 谢宇航,赵健,郭昕,等.组合预测模型在路基沉降中的应用研究[J].交通科学与工程,2019,35(4):28-33.
XIE Yuhang, ZHAO Jian, GUO Xin, et al. Research on combination forecasting model in subgrade settlement [J]. Journal of Transport Science and Engineering, 2019, 35(4): 28-33.
- [3] 刘寒冰,向一鸣,阮有兴.背景值优化的多变量灰色模型在路基沉降预测中的应用[J].岩土力学,2013,34(1):173-181.
LIU Hanbing, XIANG Yiming, RUAN Youxing. A multivariable grey model based on background value optimization and its application to subgrade settlement prediction [J]. Rock and Soil Mechanics, 2013, 34(1): 173-181.
- [4] 唐超,陈妍颖,李庶林,等.基于自适应变异粒子群算法改进OGM(1,N)及其在排土场变形预测中的应用[J].岩石力学与工程学报,2020,39(S1):3197-3205.
TANG Chao, CHEN Yanying, LI Shulin, et al. The application of improved OGM (1, N) in waste dump deformation prediction based on adaptive particle mutation swarm optimization [J]. Chinese Journal of Rock Mechanics and Engineering, 2020, 39(S1): 3197-3205.
- [5] 彭官友.桥梁施工监控的改进GM(1,1)灰色预测模型研究[J].交通科学与工程,2019,35(1):54-60.
PENG Guanyou. Application of improved GM (1, 1) grey prediction model in the bridge construction monitoring [J]. Journal of Transport Science and Engineering, 2019, 35(1): 54-60.
- [6] 彭立顺,蔡润,刘进波,等.基于遗传优化神经网络的高速公路路基沉降量预测[J].地震工程学报,2019,41(1):124-130, 207.
PENG Lishun, CAI Run, LIU Jinbo, et al. Settlement

(下转第41页)

- Journal of Highway and Transport, 2006, 19(6): 25-29.
- [10] 李灿华, 向晓东, 周溪滢. 钢渣开级配透水沥青混合料及性能研究[J]. 建筑材料学报, 2015, 18(1): 168-171.
LI Canhua, XIANG Xiaodong, ZHOU Xiying. Investigation of performance of porous open graded steel slag asphalt mixture [J]. Journal of Building Materials, 2015, 18(1): 168-171.
- [11] 叶勇, 周新星, 刘全涛. 细集料对钢渣沥青混合料黏附性的影响研究[J]. 武汉理工大学学报(交通科学与工程版), 2016, 40(3): 423-427.
- YE Yong, ZHOU Xinxing, LIU Quantao. Effect of fine aggregates on the adhesive properties of steel slag based asphalt mixtures [J]. Journal of Wuhan University of Technology (Transportation Science & Engineering), 2016, 40(3): 423-427.
- [12] 温金保. 钢渣的机械力化学效应研究[D]. 南京: 南京工业大学, 2003.
WEN Jinbao. The study of mechano-chemical effect for the steel slag [D]. Nanjing: Nanjing Tech University, 2003.
(责任编辑: 刘国奇; 校对: 欧兆虎; 英文编辑: 李岳林)

(上接第34页)

- prediction of highway subgrades based on genetic optimization neural network [J]. China Earthquake Engineering Journal, 2019, 41(1): 124-130, 207.
- [7] 李建新, 刘小生, 肖钢, 等. 基于PSR-WSVM模型的边坡位移预测[J]. 大地测量与地球动力学, 2020, 40(6): 577-580.
LI Jianxin, LIU Xiaosheng, XIAO Gang, et al. Slope displacement prediction based on PSR-WSVM model [J]. Journal of Geodesy and Geodynamics, 2020, 40(6): 577-580.
- [8] 岳振华, 沈涛, 毛曦, 等. 循环神经网络的地面沉降预测方法[J]. 测绘科学, 2020, 45(12): 145-152.
YUE Zhenhua, SHEN Tao, MAO Xi, et al. Study on prediction method of land subsidence based on recurrent neural network [J]. Science of Surveying and Mapping, 2020, 45(12): 145-152.
- [9] 刘青豪, 张永红, 邓敏, 等. 大范围地表沉降时序深度学习预测法[J]. 测绘学报, 2021, 50(3): 396-404.
LIU Qinghao, ZHANG Yonghong, DENG Min, et al. Time series prediction method of large-scale surface subsidence based on deep learning [J]. Acta Geodaetica et Cartographica Sinica, 2021, 50(3): 396-404.
- [10] CHO K, VAN MERRIENBOER B, GULCEHRE C, et al. Learning phrase representations using RNN encoder - decoder for statistical machine translation [C]// Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Doha, Qatar. Stroudsburg, PA, USA: Association for Computational Linguistics, 2014.
- [11] 杨亚波. 基于深度置信网络的边坡位移预测研究[D]. 石家庄: 石家庄铁道大学, 2020.
YANG Yabo. Study on slope displacement prediction based on depth confidence network [D]. Shijiazhuang: Shijiazhuang Tiedao University, 2020.
- [12] BOULANGER-LEWANDOWSKI N, BENGIO Y, VINCENT P. Modeling temporal dependencies in high-dimensional sequences: application to polyphonic music generation and transcription [C]// Proceedings of the 29th International Conference on International Conference on Machine Learning. 26 June 2012, Edinburgh, Scotland. New York: ACM, 2012: 1881 - 1888.
(责任编辑: 欧兆虎; 校对: 罗容; 英文编辑: 陈璐)