



计算机应用
Journal of Computer Applications
ISSN 1001-9081, CN 51-1307/TP

《计算机应用》网络首发论文

题目：基于多尺度跳跃深度长短期记忆网络的短期多变量负荷预测
作者：肖勇，郑楷洪，郑镇境，钱斌，李森，马千里
收稿日期：2020-05-31
网络首发日期：2020-10-16
引用格式：肖勇，郑楷洪，郑镇境，钱斌，李森，马千里. 基于多尺度跳跃深度长短期记忆网络的短期多变量负荷预测[J/OL]. 计算机应用.
<https://kns.cnki.net/kcms/detail/51.1307.TP.20201016.1136.006.html>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

基于多尺度跳跃深度长短期记忆网络的短期多变量负荷预测

肖勇¹, 郑楷洪¹, 郑镇境², 钱斌¹, 李森², 马千里^{2*}

(1.南方电网科学研究院, 广东 广州 510080; 2.华南理工大学 计算机科学与工程学院, 广东 广州 510006)

(*通信作者电子邮箱 qianlima@scut.edu.cn)

摘要:近年来,以循环神经网络为主体构建的预测模型在短期负荷预测中取得了优越的性能。然而,由于循环神经网络不能有效地捕捉存在于短期电力负荷数据的多尺度时序特征,因而难以进一步提升负荷预测精度。为了捕获短期电力负荷数据中的多尺度时序特征,提出了一种新颖的短期电力负荷预测模型,称为多尺度跳跃深度长短期记忆网络(MSD-LSTM)。具体来说,MSD-LSTM以长短期记忆网络(LSTM)为主体构建预测模型,能够较好地捕获长短期时序依赖,从而缓解时序过长时重要信息容易丢失的问题。进一步地,采用多层LSTM架构并且设置不同的跳跃连接数,使得每层LSTM能够捕获不同时间尺度的特征。最后,引入全连接层把各层提取到的多尺度时序特征进行融合,再利用该融合特征进行短期电力负荷预测。实验结果表明,与LSTM相比,MSD-LSTM取得了更低的均方误差,其均方误差总体下降了10%。MSD-LSTM能够更好地提取短期负荷数据中的多尺度时序特征,从而提高短期电力负荷预测精度。

关键词:短期电力负荷预测;时间序列预测;多尺度时序特征;长短期记忆网络;跳跃连接

中图分类号: TM714

文献标志码: A

Multi-scale skip deep long short-term memory network for short-term multivariate load forecasting

XIAO Yong¹, ZHENG Kaihong¹, ZHENG Zhenjing², QIAN Bin¹, LI Sen², MA Qianli^{2*}

(1. Electric Power Research Institute, China Southern Power Grid, Guangzhou Guangdong 510080, China;

2. College of Computer Science and Engineering, South China University of Technology, Guangzhou Guangdong 510006, China)

Abstract: In recent years, the short-term power load predictor built with recurrent neural network had achieved excellent performance in short-term load forecasting. However, recurrent neural network cannot effectively capture the multi-scale temporal features of short-term power load data, making it difficult to further improve the load forecasting performance. In this paper, a novel short-term power load predictor, called MSD-LSTM (Multi-scale Skip Deep Long Short-Term Memory) was proposed to capture multi-scale features of short-term power load data. Specifically, MSD-LSTM build a forecasting model with LSTM (Long Short-Term Memory), which can better capture long short-term temporal dependencies, thereby alleviating the problem that important information is easily lost when encountering the long time series. Further, a multi-layer LSTM architecture was adopted and a different skip connection was set by each layer, enabling each layer of MSD-LSTM can capture the features of different temporal scales. Finally, a fully connected layer was introduced to fuse the multi-scale temporal features extracted from each layer, and the fusion feature was used to perform the short-term load prediction. The experimental results show that compared with LSTM, MSD-LSTM achieved lower mean square error, and its mean square error decreased by 10% overall. MSD-LSTM can better capture multi-scale temporal features in short-term power load data, improving the performance of load forecasting.

Keywords: short-term power load forecasting; time series forecasting; multi-scale temporal feature; Long Short-Term Memory (LSTM) network; skip connection

0 引言

随着社会经济的发展和人民生活水平的提高,对供电质量提出了更高的要求。精确的电力负荷预测能为电量生产和供电调度提供可靠的指导,从而提升电力系统的供电质量。

收稿日期: 2020-05-31; 修回日期: 2020-07-20; 录用日期: 2020-07-22。

基金项目: 国家自然科学基金重点项目(61751205); 国家自然科学基金项目(61872148)

作者简介: 肖勇(1978—),男,湖南怀化人,高级工程师,博士,主要研究方向:电能计量管理、电能计量自动化系统(用电信息采集系统)、用电技术;郑楷洪(1991—),男,广东汕头人,工程师,硕士,主要研究方向:电能计量、电能计量自动化系统(用电信息采集系统)、用电技术;郑镇境(1996—),男,广东揭阳人,硕士研究生,主要研究方向:数据挖掘、机器学习、神经网络;钱斌(1989—),男,湖北十堰人,工程师,硕士,主要研究方向:电能计量技术;李森(1994—),男,广东茂名,硕士研究生,主要研究方向:数据挖掘、机器学习、神经网络;马千里(1980—),男,甘肃宕昌人,教授,博士,主要研究方向:数据挖掘、机器学习、神经网络。

传统的负荷预测方法有时间序列分析^[1]，回归分析^[2]和卡尔曼滤波^[3]等。随着人工智能技术的发展，神经网络^[4-5]、支持向量回归^[6]和高斯过程^[7]等智能方法被应用于电力负荷预测，其中神经网络拥有较强的非线性拟合和自学习能力，能够较好地解决传统方法非线性拟合能力差的缺陷。循环神经网络(Recurrent Neural Network, RNN)^[8]是一类拥有循环连接的神经网络，能够对时序数据进行建模。因此，在人工智能方法中使用 RNN 为主体构建的预测模型成为短期负荷预测模型的典型代表。虽然 RNN 具有时序建模的能力，但存在“梯度消失”和“梯度爆炸”的问题。为了解决该问题，Hochreiter 和 Schmidhuber 提出了长短期记忆网络(Long Short-Term Memory, LSTM)^[9]。文献[10]和文献[11]成功将 LSTM 应用到短期电力负荷预测中，提升了负荷预测精度。然而，短期电力负荷数据中存在着多尺度时序特征，其具体表现为在年、月、星期和日不同期限上均具有明显的周期性^[12]。近年来，在时间序列分类^[13]和聚类^[14]任务中，多尺度特征已经展示出能有效提升模型性能的潜力。然而，如何高效地对存在于短期电力负荷数据中的多尺度时间结构进行建模，进一步提升短期电力负荷预测的精度，仍未得到广泛的研究。

针对电力负荷数据中存在的多尺度时间结构,本文提出了一种基于多尺度跳跃深度 LSTM(Multi-scale Skip Deep Long Short-Term Memory, MSD-LSTM)的短期电力负荷预测模型。该模型使用多层 LSTM 网络架构并且每层设计不同的跳跃连接。LSTM 能够较好地对电力负荷进行建模，跳跃连接使得模型具有捕获负荷数据中多尺度时序特征的能力。实验结果证明该模型有效地提升了短期电力负荷预测的精度。

本文的主要贡献如下：

- 1) 提出了一种基于多尺度跳跃深度 LSTM 的短期电力负荷预测模型，通过使用不同的跳跃连接来捕获短期电力负荷数据中的多尺度时序结构特征，帮助模型更好的进行负荷预测；
- 2) 两个地区负荷数据的实验结果表明所提出的短期负荷预测方法可以有效提高电力负荷预测精度。

1 时间序列分析

电力负荷数据从长期来看是波动变化的，但是这种波动变化会呈现出周期性和连续性的特点。对于短期电力负荷预测而言，其在年、月、星期、日不同期限上均具有明显的周期性。

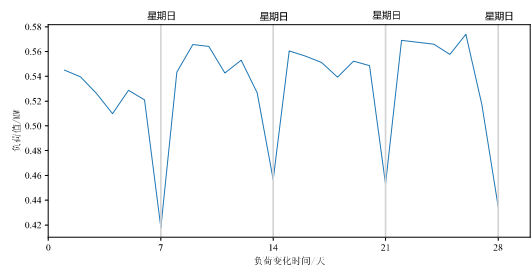


图1 某地区四周的负荷变化曲线

Fig. 1 A four-week load curve of an area

图 1 所示为某地区四个星期的电力负荷变化曲线。从图 1 可以看出，每个星期中工作日的负荷值较为平稳，此时工业和商业用电主导。在休息日，负荷从工业和商业用电主导转变为家庭用电主导，负荷值呈现下降趋势。此外，从图 1 中可以看出电力负荷值在星期这个时间尺度上呈现出明显的周期性。

通过上述分析可以发现，时间是短期电力负荷预测中重要的影响因素，这使得电力负荷数据呈现出规律波动的特点并具有明显的周期性。针对短期电力负荷数据存在的周期性变化，本文利用多层 LSTM 并且在每层设置不同的跳跃连接数，显式地建模电力负荷数据中存在的多尺度时间结构，从而更好地进行短期电力负荷预测。

2 LSTM

虽然 RNN 具有时序建模的能力，能够较好地应用于电力负荷预测任务，但在误差反向传播时，会存在“梯度消失”和“梯度爆炸”的问题。为了解决 RNN 存在的问题，Hochreiter 和 Schmidhuber 提出了 LSTM。LSTM 通过专门地设计隐含层单元来解决长期依赖问题。不同于标准的 RNN，LSTM 在隐含层的内部设计了遗忘门，输入门，输出门以及细胞状态，这使得 LSTM 能够拥有良好的长短期记忆能力。

LSTM 的隐含层单元可以看作一个细胞，细胞状态保存了长期记忆。细胞状态在每个时刻与当前输入进行少量的线性交互，从而允许网络在较长持续时间内保存信息。此外，当信息被使用，网络可以遗忘旧的状态。

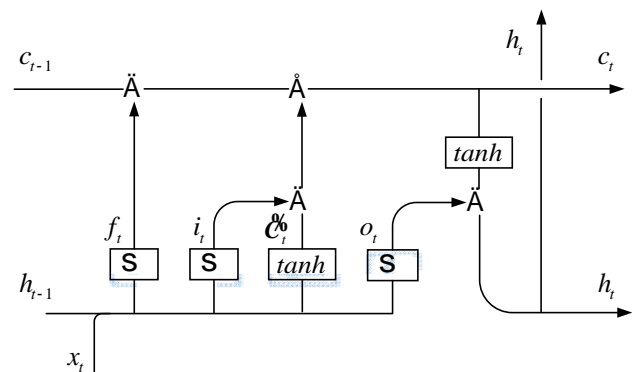


图2 LSTM隐含层内部结构图

Fig. 2 Hidden layer internal Structure of LSTM

LSTM 的隐含层内部结构图如图 2 所示, 其中 x_t 表示 t 时刻 LSTM 输入变量, h_t 和 h_{t-1} 分别表示 t 和 $t-1$ 时刻的隐含层状态, C_t 和 C_{t-1} 分别表示 t 和 $t-1$ 时刻的细胞状态, f_t 表示遗忘门的输出, i_t 表示输入门的输出, \mathcal{C}_t 表示 t 时刻用来更新细胞状态的备选信息, o_t 表示输出门的输出, s 表示 sigmoid 函数。

$$f_t = \text{sigmoid}(W_f \times [h_{t-1}, x_t] + b_f) \quad (1)$$

$$\text{sigmoid}(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

式(1)中, W_f 表示遗忘门权重矩阵, b_f 表示遗忘门偏置。根据式(1)和式(2)可知, f_t 为一个 0 到 1 的实数, 这决定了细胞状态的遗忘比例。

$$i_t = \text{sigmoid}(W_i \times [h_{t-1}, x_t] + b_i) \quad (3)$$

$$\mathcal{C}_t = \tanh(W_c \times [h_{t-1}, x_t] + b_c) \quad (4)$$

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (5)$$

同样的, 式中的 W_i 和 W_c 表示连接权重矩阵, b_i 和 b_c 表示偏置。 i_t 为 0 到 1 的实数, \mathcal{C}_t 表示 t 时刻用来更新细胞状态的备选信息, $\tanh(\cdot)$ 表示双曲正切函数。

在遗忘门中细胞状态与 f_t 相乘, 遗忘需要丢弃的信息; 在输入门中再加上 $i_t * \mathcal{C}_t$ 更新细胞状态。这可以用以下公式表示:

$$C_t = f_t \times C_{t-1} + i_t \times \mathcal{C}_t \quad (6)$$

最后, LSTM 需要确定输出什么值, 这个输出是基于经过更新的细胞状态, 具体过程可以用以下公式表示:

$$o_t = \text{sigmoid}(W_o \times [h_{t-1}, x_t] + b_o) \quad (7)$$

$$h_t = o_t \times \tanh(C_t) \quad (8)$$

式中, W_o 表示输出门权重矩阵, b_o 表示输出门偏置。简便起见, 本文将 LSTM 隐含层状态的更新过程表示为函数 $F_{\text{LSTM}}(\cdot)$, 在长短期记忆网络的前向传播过程中, 隐含层状态的更新公式为:

$$h_t = F_{\text{LSTM}}(h_{t-1}, x_t) \quad (9)$$

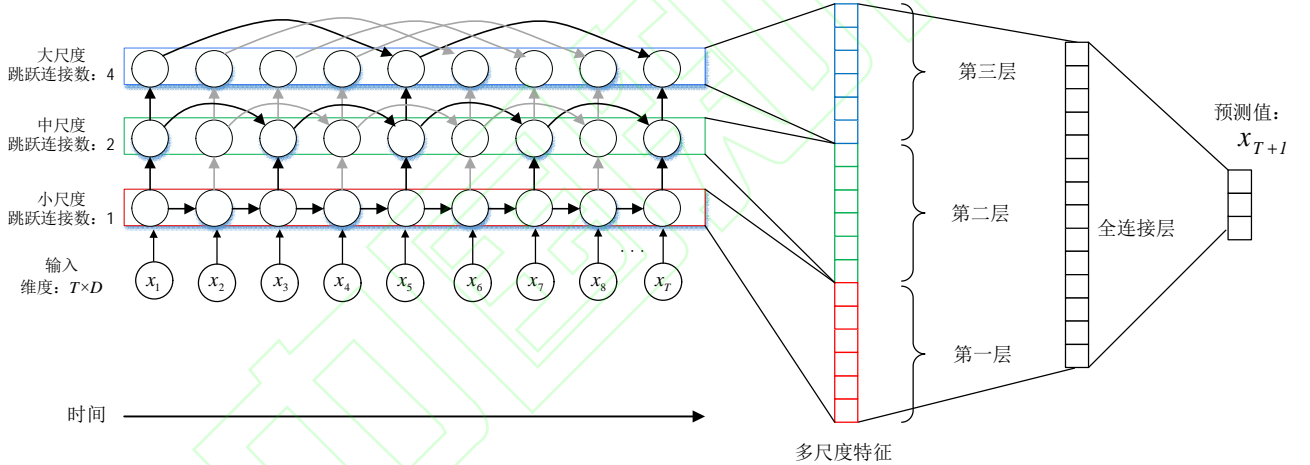


图3 多尺度跳跃深度长短期记忆网络结构图。

Fig. 3 Structure of MSD-LSTM

3 基于 MSD-LSTM 的负荷预测模型

3.1 MSD-LSTM

LSTM 虽然能够有效地解决了 RNN 存在的无法有效学习长时依赖的问题, 但是由于电力负荷数据中存在着不同尺度的时间结构特征, 而 LSTM 却不具备捕捉多尺度时间特征的能力。针对该问题, 本文对 LSTM 进行改进, 提出了一种基于多尺度跳跃深度长短期记忆网络(Multi-scale Skip Deep Long Short-Term Memory, MSD-LSTM)的短期负荷预测方法, 使其具有捕捉电力负荷数据中多尺度时序信息的能力。MSD-LSTM 通过把 LSTM 的拓扑结构设计为多层并且加入循环跳跃连接^[15], 以此改变历史信息与当前信息的交互周

期, 使得每层拥有了不同时间周期的建模能力, 显式地对电力负荷数据中的多尺度时间结构进行建模。MSD-LSTM 的每一层使用不同循环跳跃尺度的 LSTM, 即每个时刻的隐含层状态更新时依赖的不是上一个时刻的隐含层状态, 而是依赖多个时刻之前的状态。如图 3 所示, 第一层的跳跃连接数为 1, 则第一层为传统的 LSTM; 第二层的跳跃连接数为 2, 则隐含层状态更新时依赖的是往前两个时刻的隐含层状态; 第三层的跳跃连接数为 4, 则隐含层状态更新时依赖的是往前四个时刻的隐含层状态。加入循环跳跃连接后的 LSTM 的更新过程可以表示为:

$$h_t^i = F_{\text{LSTM}}(h_{t-k}^i, x_t^i) \quad (10)$$

$$y_t^i = \tanh(V^i \times h_t^i) \quad (11)$$

式中上标 i 表示层数, 后续模型部分的所有上标皆为该含义。

在网络训练的前向传播过程中, 第一层 LSTM 的输入为预处理后的电力负荷数据以及天气数据组成的多维变量序列, 如图 3 中输入数据为 1 到 T 时刻的负荷数据, 维度为 $T \times D$, T 表示输入的时刻数, D 表示每个时刻输入的向量的维度; 输出为 $T+1$ 时刻的负荷预测值。天气数据帮助模型更好地进行负荷预测。其余层的输入为上一层 LSTM 的输出, 最后把每一层最后一个时刻的隐含层状态进行拼接后输入到全连接层形成融合特征, 利用该融合特征进行未来负荷值的预测。多层不同跳跃尺度的 LSTM 前向传播过程可以用下列公式表示:

$$\begin{aligned} h_t^1 &= F_{\text{LSTM}}(h_{t-k_1}^1, x_t) \\ y_t^1 &= \tanh(V^1 \mathcal{H}_t^1) \\ h_t^2 &= F_{\text{LSTM}}(h_{t-k_2}^2, y_t^1) \\ y_t^2 &= \tanh(V^2 \mathcal{H}_t^2) \\ &\vdots \\ h_t^L &= F_{\text{LSTM}}(h_{t-k_L}^L, y_t^{L-1}) \\ y_t^L &= \tanh(V^L \mathcal{H}_t^L) \end{aligned} \quad (12)$$

MSD-LSTM 的网络结构图如图 3 所示, 其中方框内的圆形表示图 2 的 LSTM 单元, 圆形左边的箭头表示接收上一个时刻的隐含层状态; 右边的箭头当前时刻的隐含层状态输入到下一个时刻的 LSTM 单元; 下方的箭头表示当前时刻的输入; 上方的箭头表示当前时刻的 LSTM 单元的输出。使用前 T 个时刻的数据来预测 $T+1$ 时刻的负荷值, 每个时刻的输入为一个 6 维向量, 即 D 为 6, 包含负荷数据以及天气数据。网络层数设为 3 层, 每一层的跳跃连接数分别为 1, 2 和 4。最后, 通过全连接层将每层 LSTM 第 T 个时刻的隐含层状态进行融合得到融合特征, 利用该融合特征去预测下一时刻的负荷值。

MSD-LSTM 通过给每层 LSTM 设计不同的跳跃连接数, 使得每层拥有不同时间周期的建模能力。本文研究的负荷预测是短期电力负荷预测, 其在年、月、星期、日不同期限上均具有明显的周期性。因此, MSD-LSTM 第一层的跳跃连接数设为 1, 捕捉日依赖关系; 第二层的跳跃连接数设为 7, 捕捉周依赖关系; 第三层的跳跃连接数设为 30, 捕捉月依赖关系。最后通过引入全连接层, 把每层提取到的不同尺度信息进行融合, 即将每层 T 时刻的隐含层状态拼接后作为全连接层的输入, 实现多尺度结构信息的融合, 并把全连接层的输出作为预测值, 即 $T+1$ 时刻的负荷值。全连接层可以使用以下公式表示:

$$x_{T+1} = W_{fc} \mathcal{H}_T + b \quad (13)$$

式中, W_{fc} 为全连接层权重, b 表示偏置, x_{T+1} 为全连接层的输出, \mathcal{H}_T 为各层 LSTM 最后一个时刻隐含层状态的拼接, 可以用以下公式表示:

$$\mathcal{H}_T = \text{concat}(\mathcal{H}_T^1, h_T^2, L, h_T^L) \quad (14)$$

3.2 基于 MSD-LSTM 的短期负荷预测算法流程

步骤 1: 构建时间序列数据进行训练和测试。对电力负荷数据以及天气数据进行归一化, 采取一步预测的方法并对归一化后的数据做窗口滑动处理得到 N 条长度为 T 的负荷序列样本集, 即以前 T 个时间点的数据作为“特征”, 下一个时间点的数据作为“标签”。将样本集按 7: 3 的比例划分为训练集和测试集, 即前 $0.7 \times N$ 条序列作为训练集, 后 $0.3 \times N$ 条序列作为测试集。

步骤 2: 逐层构建 LSTM, 每一层设置不同的跳跃连接以建模多尺度时序依赖关系。

步骤 3: 构建全连接层将各层 LSTM 的依赖关系进行融合并把全连接层的输出作为预测值形成多尺度跳跃深度 LSTM 模型。

步骤 4: 训练步骤 3 构建的模型, 在网络训练的前向传播过程中, 第一层 LSTM 输入为经步骤 1 处理过的训练数据集, 其余层的输入为上一层 LSTM 的输出, 把每一层最后一个时刻的隐含层状态进行拼接后输入到全连接层形成融合特征, 利用该融合特征进行未来负荷值的预测。在网络训练的反向传播过程中, 使用基于时间的反向传播算法 (Back Propagation Through Time, BPTT)^[16] 来最小化损失函数, 即均方误差, 从而优化模型的参数。

步骤 5: 使用训练后的模型对经步骤 1 处理过的电力负荷测试数据集进行预测, 计算预测值和真实值之间的均方误差。

4 实验分析与比较

4.1 数据来源

本文研究的是带有天气数据的短期电力负荷预测, 使用 2012 年 1 月 1 日到 2014 年 12 月 31 日两个地区真实负荷数据, 数据来源于 2016 年电工数学建模竞赛 A 题。此外, 这个数据集还提供最高温度、最低温度、平均温度、相对湿度和降雨量五个天气数据。本文利用这五个天气数据来帮助负荷预测, 即将负荷数据以及五个天气数据拼接为 6 维向量。

本文分别抽取两个地区每天 6 点、12 点和 18 点的负荷值进行负荷预测, 因此总的样本集有 6 个, 每个样本集中有 1096 个时间点, 每个时间点为一个 6 维的向量, 包含负荷及天气数据。负荷和天气数据拥有不同的数量级和量纲。为了去除这些影响, 需要对输入量进行归一化处理, 将其转化为无量纲的纯数值。本文采取离差标准化的归一化方式。每个

样本集原始数据表示为 $s = [s_1, s_2, \mathbf{L}, s_n]^T$ 其中 $s_j = \hat{s}_j^1, s_j^2, \mathbf{L}, s_j^6 \hat{u}^T$, 对原始数据进行如下归一化:

$$e_j^d = \frac{s_j^d - \min(s_{1:n}^d)}{\max(s_{1:n}^d) - \min(s_{1:n}^d)} \quad (15)$$

得到归一化的样本集 $e = [e_1, e_2, \mathbf{L}, e_n]^T$ 。

4.2 实验设置

这部分展示了 MSD-LSTM 应用在短期电力负荷预测上的实验设置。本文将 MSD-LSTM 与多层长短期记忆网络(Multi-layer Long Short-Term Memory, MLSTM), 单层长短期记忆网络(Single-layer Long Short-Term Memory, SLSTM), 支持向量回归(Support Vector Regression, SVR)^[17]和高斯过程(Gaussian Process, GP)^[18]进行了对比实验。使用过去 35 天的负荷值和天气数据来预测下一天的负荷值, 则对样本集做窗口滑动处理得到 1062 条长度为 35 的序列, 即样本。构建时间序列数据进行训练和测试时, 将 1062 个样本集按 7: 3 的比例划分为训练集和测试集。MSD-LSTM 的层数设为 3, 每层的跳跃尺度分别设为 1、7、30, 相应的每层分别对时间尺度以天、周、月为单位的电力负荷数据时序依赖关系进行捕捉。为了检验 MSD-LSTM 对电力负荷数据中的多尺度时间结构的捕捉能力, 本文设置了 MLSTM 和 SLSTM 两个对比实验。MLSTM 的层数设为 3, 每一层的跳跃尺度设为 1, 以检验多尺度时间结构对预测性能的提升; SLSTM 的层数设为 1, 跳跃尺度设为 1, 以检验深层网络结构对预测性能的提升。LSTM 的隐含层单元数固定为 48。为了保证对比实验的公平, 本文保持上述三种模型在每个样本集上所使用的隐含层单元

数一致。SVR 和 GP 同样使用过去 35 天的负荷值和天气数据来预测下一天的负荷值。每天的数据包含负荷值以及 5 个天气数据, 因此本文将 35 天的数据拼接成一个 35×6 的向量作为模型的输入, 并预测下一天的负荷值。此外, SVR 和 GP 都是基于核函数的模型, 在本文中 SVR 和 GP 的核函数都选取高斯核函数。

MSD-LSTM、MLSTM 和 SLSTM 训练时使用 mini-batch 技术, batch size 设置为 32, 即训练时每次迭代输入 32 个样本并计算损失函数, 根据损失函数反向传播更新网络的参数。训练时使用 mini-batch 技术, 可以使网络参数更新更快, 同时有利于网络参数更加鲁棒地收敛。此外, 为了防止过拟合, 在全连接层使用 Dropout, Dropout 率设为 0.05。

4.3 实验结果

4.3.1 对比实验

使用 MSD-LSTM 与 MLSTM, SLSTM, SVR 和 GP 分别进行实验并对比。表 1 为各模型在短期电力负荷上的预测结果, 采用均方误差(Mean Squared Error, MSE)作为评判标准:

$$MSE = \frac{1}{M} \sum_{i=1}^M (y^i - y_{pred}^i)^2 \quad (16)$$

其中 y^i 和 y_{pred}^i 分别表示测试集中对应第 i 个样本的真实值和预测值, M 表示测试集样本个数。

表1 不同方法的负荷预测误差(MSE)对比

Tab. 1 Comparison of load forecasting error (MSE) of different methods

样本集编号	MSD-LSTM	MLSTM	SLSTM	SVR	GP
1	5.20E-03	6.07E-03	6.13E-03	1.05E-02	1.55E-02
2	9.13E-03	1.04E-02	1.03E-02	1.44E-02	2.06E-02
3	9.38E-03	1.11E-02	1.10E-02	1.43E-02	1.96E-02
4	2.43E-03	3.02E-03	3.16E-03	8.90E-03	4.61E-03
5	5.20E-03	6.01E-03	6.14E-03	1.00E-02	6.22E-03
6	5.03E-03	5.58E-03	5.53E-03	1.05E-02	5.46E-03

从表 1 可以看出, MSD-LSTM 在六个样本集中的短期负荷预测都取得了最好的预测精度。一方面, 与 SVR 和 GP 对比, MSD-LSTM 是以 LSTM 为主体构建的预测模型, 能够较好地处理序列建模任务, 而 SVR 和 GP 并未对存在于电力负荷数据中的时序依赖关系进行建模, 因此 MSD-LSTM 的预测性能要明显优于这两种方法。另一方面, 与 SLSTM 对比, MSD-LSTM 由三层 LSTM 构成, 能够更好地建模电力负荷数据中的动态特性, 取得了更好的预测精度, 表明深层网络结构有利于预测性能的提升。与 MLSTM 对比,

MSD-LSTM 通过循环跳跃连接对短期电力负荷中的日、周和月的时序依赖关系进行建模, 增强了 LSTM 对多尺度时序结构建模的能力, 取得了更好的预测精度, 表明多尺度时间结构能够提升预测性能。因此, MSD-LSTM 通过结合深层网络结构建模动态特性的能力以及多尺度跳跃连接提取多尺度时间结构信息的能力, 有效地提升了短期电力负荷预测精度。

更进一步的, 本文从 6 个样本集中随机抽取 1 个样本集, 使用上述各种模型对其进行负荷预测, 并对比了各种模型的绝对误差。绝对误差越接近 0 说明预测精度越高, 越远离 0

说明预测精度越低。图 4 展示了 MSD-LSTM、SLSTM、MLSTM、SVR 和 GP 五种模型预测的负荷值与实际负荷值的绝对误差曲线。对比五张绝对误差曲线图，可以看出

MSD-LSTM 能够取得最好的预测精度，具体表现为取得了最低的绝对误差平均值 0.069。

表2 不同超参数设置的负荷预测误差(MSE)对比

Tab. 2 Comparison of load forecasting error (MSE) of different hyperparameter settings

实验组	层数	跳跃数	步长	样本集编号					
				1	2	3	4	5	6
1	2	1, 2	35	5.70E-03	9.96E-03	1.11E-02	2.53E-03	5.80E-03	5.60E-03
2	2	1, 7	35	5.46E-03	9.21E-03	9.62E-03	2.52E-03	5.87E-03	5.44E-03
3	2	1, 7	25	5.23E-03	9.12E-03	1.04E-02	2.53E-03	5.49E-03	5.21E-03
4	3	1, 2, 4	35	5.59E-03	1.00E-02	1.11E-02	2.47E-03	5.56E-03	5.21E-03
5	3	1, 7, 14	35	5.53E-03	9.78E-03	9.69E-03	2.55E-03	5.31E-03	5.15E-03
6	3	1, 7, 30	45	5.38E-03	9.07E-03	1.04E-02	2.76E-03	5.38E-03	5.17E-03
7	3	1, 7, 30	35	5.20E-03	9.13E-03	9.38E-03	2.43E-03	5.20E-03	5.03E-03

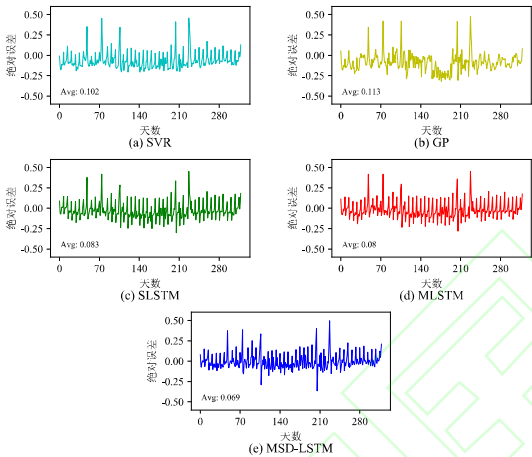


图4 不同方法的绝对误差曲线

Fig. 4 The absolute error curve of different methods

4.3.2 超参数探究

为了探究层数、跳跃数以及步长（即使用过去多少天的数据来进行下一天的负荷预测）对 MSD-LSTM 性能的影响，本文设置了 7 组使用不同超参数的对比实验，其在 6 个样本集上的结果如表 2 所示。首先，对比实验 1 和实验 4 的结果，实验 4 取得了更低 MSE，这表明深层网络有利于提升短期负荷预测的精度。对比实验 2 和实验 7 可以得到同样的结果。进一步的，对比实验 4、5 和 7 的结果，实验 7 取得了最低 MSE。这表明跳跃数会影响模型的性能，同时根据数据实际的多尺度时间结构来选择合适的跳跃数有利于提升模型的性能。最后，对比实验 2、3 以及实验 6、7 的结果，可以发现具有更短步长的模型可以取得更好的结果。因此，有效地使用 MSD-LSTM 需要综合考虑多个超参数的影响。

5 结语

本文针对 LSTM 无法有效建模电力负荷数据中存在的多尺度时间结构，提出了 MSD-LSTM。MSD-LSTM 将 LSTM 的拓扑结构设计为多层并在每层中设置不同的跳跃连接，显式地对短期电力负荷数据中的多尺度时间结构进行建模。通过两个地区负荷数据的实验分析，证明了 MSD-LSTM 能够有效地捕捉短期电力负荷数据中的多尺度时间结构，从而提升负荷预测精度。

MSD-LSTM 采取的跳跃连接数是手工设置的，需要依靠先验知识，无法自适应地学习电力负荷数据的多尺度结构。未来的工作会研究如何动态自适应地学习跳跃连接，更好地根据电力负荷数据来学习多尺度结构并对其进行建模，进一步提高电力负荷预测的精度。

参考文献

[1] 崔和瑞,彭旭.基于 ARIMAX 模型的夏季短期电力负荷预测[J].电力系统保护与控制,2015,43(04):108-114. (CUI H R, PENG X. Summer Short-Term Load Forecasting Based on ARIMAX Model[J]. Power System Protection and Control, 2015, 43(04): 108-114.)

[2] DUDEK G. Pattern-based local linear regression models for short-term load forecasting[J]. Electric Power Systems Research, 2016, 130: 139-147.

[3] TAKEDA H, TAMURA Y, SATO S. Using the ensemble Kalman filter for electricity load forecasting and analysis[J]. Energy, 2016, 104: 184-198.

[4] 吴云,雷建文,鲍丽山等.基于改进灰色关联分析与蝙蝠优化神经网络的短期负荷预测[J].电力系统自动化,2018,42(20):73-78. (WU Y, LEI J W, BAO L S, et al. Short-term Load Forecasting Based on Improved Grey Relational Analysis and Neural Network Optimized by Bat Algorithm[J]. Automation of Electric Power Systems, 2018, 42(20): 73-78.)

[5] 罗育辉,蔡延光,戚远航等.基于最大偏差相似性准则的 BP 神经网络短期电力负荷预测算法[J].计算机应用研究,2019,36(11): 3269-3273. (LUO Y H, CAI Y G, QI Y H, et al. Short-term power load forecasting algorithm based on maximum deviation similarity criterion BP neural network[J]. Application Research of Computers, 2019, 36(11): 3269-3273.)

[6] FAN G F, PENG L L, HONG W C, et al. Electric load forecasting by the SVR model with differential empirical mode decomposition and auto regression[J]. Neurocomputing, 2016, 173: 958-970.

- [7] 任利强,张立民,王海鹏等.基于 IPSO-GPR 的短期负荷区间预测[J]. 计算机工程与设计,2019,40(10):3002-3008. (REN L Q, ZHANG L M, WANG H P, et al. Short-term power load interval forecasting based on improved particle swarm optimization and Gaussian process regression[J]. Computer Engineering and Design, 2019, 40(10): 3002-3008.)
- [8] ELMAN J L. Finding Structure in Time[J]. Cognitive Science, 1990, 14(2): 179-211.
- [9] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [10] ZHENG J, XU C, ZHANG Z, et al. Electric load forecasting in smart grids using long-short-term-memory based recurrent neural network[C]// Proceedings of the 2017 51st Annual Conference on Information Sciences and Systems. Piscataway, NJ: IEEE, 2017: 1-6.
- [11] 李鹏,何帅,韩鹏飞等.基于长短期记忆的实时电价条件下智能电网短期负荷预测[J].电网技术,2018,42(12):4045-4052. (LI P, HE S, HAN P F, et al. Short-Term Load Forecasting of Smart Grid Based on Long-Short-Term Memory Recurrent Neural Networks in Condition of Real-Time Electricity Price[J]. Power System Technology, 2018, 42(12): 4045-4052.)
- [12] 康重庆,夏清,张伯明.电力系统负荷预测研究综述与发展方向的探讨[J].电力系统自动化,2004,28(17):1-11. (KANG C Q, XIA Q, ZHANG B M. Review of power system load forecasting and its development[J]. Automation of Electric Power Systems, 2004, 28(17): 1-11.)
- [13] MA Q, CHEN E, LIN Z, et al. Convolutional multitimescale echo state network[J]. IEEE Transactions on Cybernetics, 2019, doi: 10.1109/TCYB.2019.2919648.
- [14] MA Q, ZHENG J, LI S, et al. Learning Representations for Time Series Clustering[C]// Proceedings of the 2019 33rd Annual Conference on Neural Information Processing Systems. New York, NY: Curran Associates, 2019: 3781-3791.
- [15] CHANG S, ZHANG Y, HAN W, et al. Dilated Recurrent Neural Networks[C]// Proceedings of the 2017 31st Annual Conference on Neural Information Processing Systems. New York, NY: Curran Associates, 2017: 77-87.
- [16] LILLICRAP T P, Santoro A. Backpropagation through time and the brain[J]. Current opinion in neurobiology, 2019, 55: 82-89.
- [17] SMOLA A J, SCHLKOPF B. A tutorial on support vector regression[J]. Stats and Computing, 2004, 14(3): 199-222.
- [18] RASMUSSEN C E. Gaussian Processes in Machine Learning[C]// Proceedings of the Summer School on Machine Learning. Berlin: Springer, 2003: 63-71.

This work is partially supported by the Key Projects of National Natural Science Foundation of China (61751205), and the National Natural Science Foundation of China (61872148).

XIAO Yong, born in 1978, Ph. D., senior engineer. His research interests include electric energy measurement management, electric energy measurement automation system, electricity technology.

ZHENG Kaihong, born in 1991, M. S., engineer. His research interests include electric energy measurement, electric energy measurement automation system, electricity technology.

ZHENG Zhenjing, born in 1996, M. S. candidate. His research interests include data mining, machine learning, neural network.

QIAN Bin, born in 1989, M. S., engineer. His research interests include electric energy measurement.

LI Sen, born in 1994, M. S. candidate. His research interests include data mining, machine learning, neural network.

MA Qianli, born in 1980, Ph. D., professor. His research interests include data mining, machine learning, neural network.