

基于 LSTM 的 PM_{2.5} 浓度预测模型

段大高, 赵振东, 梁少虎, 杨伟杰, 韩忠明

(北京工商大学 计算机与信息工程学院, 北京 100048)

摘要: 随着近年雾霾天气的频繁出现, 空气质量开始越来越受到公众关注; PM_{2.5} 浓度指数是判断空气质量的重要指标, 如何根据历史数据有效地预测空气中 PM_{2.5} 浓度, 具有很高的应用价值; 分析以往空气质量数据表明, PM_{2.5} 浓度有明显的非线性和不确定性波动, 很难用传统机器学习算法有效地预测; 文章基于 LSTM 循环神经网络, 依据过去 20 小时采集的空气数据, 预测未来 5 小时的 PM_{2.5} 浓度指数; 实验结果表明, LSTM 可以有效地捕获空气质量的时序特征, 较准确预测出未来时刻的 PM_{2.5} 浓度指数。

关键词: PM_{2.5}; LSTM 循环神经网络; 时序特征

Research on PM_{2.5} concentration prediction based on LSTM

Duan Dagao, Zhao Zhendong, Liang Shaohu, Yang Weijie, Han Zhongming

(School of Computer and Information Engineering, Beijing Technology and Business University, Beijing 100048, China)

Abstract: With the frequent occurrence of smoggy weather in recent years, air quality has begun to receive more and more public attention. The PM_{2.5} index is an important indicator for judging air quality. How to effectively predict the PM_{2.5} concentration in the air based on historical data, which has application value. And analysis of previous air quality data show that PM_{2.5} concentration has obvious nonlinearity and uncertain volatility, which is difficult to predict effectively with traditional machine learning algorithms. Based on the LSTM (long short-term memory) recurrent neural network (RNN), this paper predicts the PM_{2.5} concentration for the next 5 hours based on the air data collected over the past 20 hours. Experiments show that LSTM can effectively capture the feature of time series for air quality and accurately predict the PM_{2.5} concentration in the future.

Keywords: PM_{2.5}; LSTM recurrent neural network; feature of time series

0 引言

随着近几年雾霾天气在全国范围内的频繁出现, 细微颗粒物 (Fine Particulate Matter, PM_{2.5}) 受到了公众的广泛关注。PM_{2.5} 通常指环境空气中空气动力学当量直径小于等于 2.5 微米的颗粒物。PM_{2.5} 能较长时间悬浮于空气中, 其在空气中含量浓度越高, 代表空气污染越严重。与较粗的环境空气颗粒物相比, PM_{2.5} 粒径小, 活性强, 极易附带有毒、有害物质 (例如, 重金属、微生物等), 且在大气中的停留时间长、输送距离远, 因而对人体健康和大气环境质量的影响更大。如何有效的对空气 PM_{2.5} 浓度进行准确的预测和预报, 对于保护公众身体健康, 环境治理具有重要的意义。由于 PM_{2.5} 浓度受到多方面因素的影响, 呈现出明显的不规则和不确定性波动, 很难对其进行

有效的数学建模。

近年来深度学习技术^[1]在人工智能领域取得显著成就。神经网络可以对数据抽象特征进行提取, 并且拥有强大的对高维数据进行拟合的能力, 可以有效地根据 PM_{2.5} 历史数据, 构建预测模型。本文基于深度学习中 LSTM^[2] (Long Short-Term Memory) 循环神经网络, 依据过去 20 小时采集的空气数据, 预测未来 5 个小时 PM_{2.5} 浓度指数。实验数据表明本文提出的算法模型达到了良好的效果, 可以有有效的预测出未来 5 小时内空气中 PM_{2.5} 浓度值。

1 相关工作

最近几年, 学术界对 PM_{2.5} 浓度预测进行了许多研究。文献 [3] 基于非线性回归模型建立预测模型, 并设计出一个基于逆风 PM_{2.5} 浓度的附加参数, 用于增强预测模型的表现, 取得了良好的效果。文献 [4] 首次提出了一种基于数据预处理和分析的混合 EEMD-GRNN (集合经验模态分解-通用回归神经网络) 模型, 用于提前一天预测 PM_{2.5} 浓度。文献 [5] 提出优化神经网络的大气 PM_{2.5} 污染指数预测方法, 利用主成分分析法对大气 PM_{2.5} 污染指数的各种影响因素进行分析, 保留影响因素的主要特征成分, 并作为神经网络的训练样本, 利用遗传算法进行 BP 神经网络结构参数的寻优, 并利用最优参数构建 BP 神经网络的预测模型, 获得准确地预测结果。文献 [6] 使用贝叶斯 kriged 卡尔曼滤波模型对 PM_{2.5} 时空过程进行短期预

收稿日期: 2018-10-09; 修回日期: 2018-10-25。

基金项目: 国家自然科学基金面上项目 (61170112); 北京市自然科学基金项目 (4172016); 北京市教委科研计划面上项目 (KM201710011006)。

作者简介: 段大高 (1976-), 男, 湖南邵阳, 博士, 副教授, 主要从事多媒体处理、数据挖掘等方向的研究。

杨伟杰 (1980-) 女, 山东潍坊人, 博士, 讲师。主要从事自然语言处理、文本数据挖掘方向的研究。

韩忠明 (1972-), 男, 山西太原人, 博士, 教授, 主要从事数据挖掘方向的研究。

测, 利用 Kriging 方法建立模型的空间预测并使用马尔可夫链蒙特卡罗技术实现, 在时间和空间上获得良好的预测效果。文献 [7] 提出了一种基于地理的模型, 使用 MLP 提前三天预测 SO₂, CO 和 PM₁₀ 的日平均浓度, 其采用了 3 种地理模型: 单站点邻域模型, 双站点邻域模型和基于距离的模型。实验结果表明, 基于地理的模型优于普通模型, 特别是基于距离的模型。如果在地理模型中增加更多的气象变量, 预计仍有很大的改进空间。文献 [8] 提出了一种新的混合 ARIMA-ANN 模型, 以提高 PM₁₀ 预报精度, 取得了较好的预测精度。

尽管之前许多学者在 PM_{2.5} 和其他空气质量预测算法上取得了不少突破性进展, 但是大部分现有算法模型都需要复杂的数据处理和设置外部参数条件, 而这一过程需要很强的专业知识, 增加了模型的实现复杂程度, 同时也降低了其易用性。本文提出的基于 LSTM 循环神经网络 PM_{2.5} 预测算法模型, 能够有效地根据历史数据预测出未来时刻 PM_{2.5} 浓度值, 同时模型具较低的实现复杂度和开发成本。

2 基于 LSTM 的 PM_{2.5} 预测算法

2.1 RNN 神经网络

原始的神经网络一般都是全连接网络, 而且非相邻的网络之间没有连接, 没有办法有效处理不定长的时序数据。RNN (Recurrent Neural Network) 是一类用于处理序列数据的神经网络, 为了更有效的处理时序数据, RNN 将网络的隐藏层相连接, 可以理解为多个具有相同结构和参数的前向神经网络的循环堆积。循环的次数和输入序列的长度一致, 且要求在序列中毗邻状态的对应网络的隐层之间互联。RNN 网络结构示意图如图 1 所示。

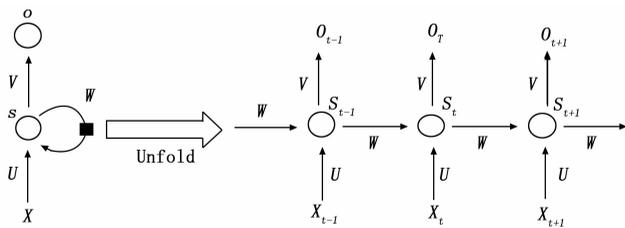


图 1 RNN 结构图

在 RNN 中网络每一个时刻除了要接受输入层的参数以外, 还要接收自身网络上一时刻隐层输出。RNN 的隐层控制着序列数据的信息传递。不同的时刻之前连接的权重代表着过去时刻信息对当前时刻的影响。

RNN 网络的输入记为 $\{x_0, x_1, \dots, x_t, x_{t+1}\} \dots$, 输出集记为 $\{y_0, y_1, \dots, y_t, y_{t+1}\} \dots$, 隐藏层的输出记为 $\{s_0, s_1, \dots, s_t, s_{t+1}\} \dots$ 。RNN 能够处理时序信息隐层单元起了关键作用。数据信息的流向是从输入层到输出层, 并跟随隐藏层的状态传递下去。具体计算过程如公式 (1) (2) (3) 所示。

$$s_t = \sigma(Ux_t + Ws_{t-1} + b_t) \quad (1)$$

$$o_t = Vh_t + b_o \quad (2)$$

$$y_t = \text{softmax}(o_t) \quad (3)$$

在上式中 U, W, V 分别是网络输入层, 隐藏层, 和输出层权重参数, b_i 和 b_o 分别是输入层偏置和输出层偏置参数, σ 是激活函数, 一般会选择 \tanh 函数。

理论上 RNN 可以处理任何长度序列数据, 但是当处理序列过长时会导致历史时刻信息的影响削弱甚至消失, 称之为梯度消失。对 RNN 的改进主要是集中在隐层神经网络的结构做优化。如 LSTM 和 GRU^[9] (Gated Recurrent Unit) 在隐层网络中增加了门控制单元, 使得网络可以保留住历史时刻比较重要的信息。

2.2 LSTM 神经网络

RNN 的关键作用是可以依据隐藏层之前的转态传递处理序列问题, 即做当前时刻的任务是可以考虑到过去时刻的影响。但是 RNN 处理较长时序数据时会很难训练, 表现极差。LSTM 是长短期记忆网络, 通过相应的网络结构改变来避免长期依赖问题。在 RNN 中通过隐藏层连接重复同一个网络模块。LSTM 同样延续了与 RNN 同样的重复连接结构, 不同的是在 LSTM 网络中增加细胞状态 c_t 。 c_t 也会随着不同的时刻传递下去, 细胞转态 c_t 代表了长期记忆。LSTM 的关键就是细胞状态, 在整个序列的运算过程中只有少量的线性交互, 所以可以有效保存过去较长时刻的信息。

LSTM 的细胞结构如图 2 所示, 当接收到前一个时刻隐藏层输出 h_{t-1} 和当前时刻输入 x_t , 首先会用遗忘门决定细胞状态丢弃不重要的信息。遗忘门运算如公式 (4) 所示:

$$f_t = \sigma(W_{xf} x_t + W_{hf} h_{t-1} + W_{cf} c_{t-1} + b_f) \quad (4)$$

在上式中 $W_{xf}, W_{hf}, W_{cf}, b_f$ 代表遗忘门的网络参数。 σ 代表 sigmoid 激活函数, f_t 是遗忘门输出向量, 其中向量中的每一个元素都在 (0, 1) 范围内。代表着当前时刻细胞状态 c_t 中每个元素的重要程度。

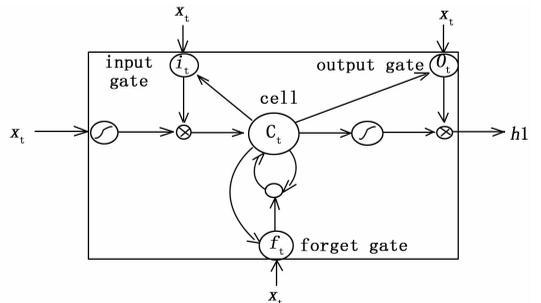


图 2 LSTM 神经元示意图

LSTM 还要确定当前时刻, 细胞状态哪些位置的信息需要更新和确定更新后的值。这个决定通过输入门来完成。输入门的运算过程如公式 (5) (6) 所示:

$$i_t = \sigma(W_{xi} x_t + W_{hi} h_{t-1} + W_{ci} c_{t-1} + b_i) \quad (5)$$

$$\tilde{c}_t = \tanh(W_{xc} x_t + W_{hc} h_{t-1} + b_c) \quad (6)$$

在上式 (5) 中 W_{xi}, W_{hi}, W_{ci} 和 b_i 都是输入门的参数, 代表 sigmoid 激活函数。公式 (6) 中函数处理输入数据 x_t 和

上个时刻的隐层输出 h_{t-1} 创建一个新候选细胞状态 \tilde{c}_t 。

当遗忘门和输入门运算完成, LSTM 会更新当前时刻的细胞状态, 更新细胞状态的计算过程如公式 (7) 所示:

$$c_t = f_t c_{t-1} + i_t \tilde{c}_t \quad (7)$$

式中, f_t 是遗忘门的输出, 代表着哪些信息需要遗忘, 哪些信息需要保留。

LSTM 最后会计算当前需要确定什么样的信息需要输出, 这个任务主要靠输出门根据当前的细胞状态 c_t 来完成。公式 (8) 计算确定哪些部分的细胞状态需要输出, 公式 (9) 把细胞状态通过 \tanh 函数处理, 得到一个 $(-1, 1)$ 范围的值, 并将其和 o_t 相乘得到当前时刻最终的隐藏层输出。

$$o_t = \sigma(W_{xo} x_t + W_{ho} h_{t-1} + b_o) \quad (8)$$

$$h_t = o_t \tanh(c_t) \quad (9)$$

式中, W_{xo} , W_{ho} 和 b_o 是输出门的参数, h_t 是当前时刻的隐层输出。

LSTM 在 RNN 的基础上增加了三个门控制, 分别是遗忘门, 输入门和输出门。遗忘门和输入门负责对细胞状态的 c_t 的更新。在网络中 c_t 会保留住序列数据的重要信息并可以传递较长的时刻, 能够有效的缓解长期依赖。LSTM 目前已被广泛的应用于序列标注, 机器翻译和语音识别等自然语言处理等任务中。

2.3 PM2.5 预测模型

本文采用 LSTM 神经网络实现, 算法模型根据当前时刻最近 20 个小时的历史空气质量观测数据预测未来 5 个小时 PM2.5 浓度值。这是个典型的时序预测问题, LSTM 能够有效的处理时序数据。算法网络模型的整体结构图如下图 3 所示。

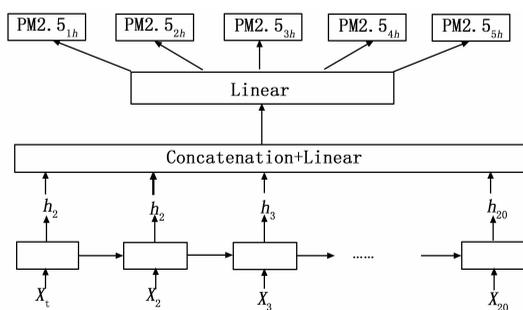


图 3 算法模型结构图

在模型中 $x_1, x_2 \dots x_{20}$ 代表某个观测地点采集的最近 20 个小时历史空气数据。 $h_1, h_2 \dots h_{20}$ 代表每个时刻输出隐状态向量。对于第 t 时刻输入数据 $x_t = \{PM2.5, PM2.5_{24h}, PM10, PM10_{24h}, AQI\}$ 。 x_t 为 5 维的向量数据, 其中 $PM2.5$ 是 t 时刻的 $PM2.5$ 测量值, $PM2.5_{24h}$ 是 t 时刻 $PM2.5$ 最近的 24 小时测量均值, $PM10$ 是时刻的 $PM10$ 测量值, $PM10_{24h}$ 是时刻 $PM10$ 最近的 24 小时测量均值。 AQI (Air Quality Index) 代表时刻的空气污染指数。模型会依据 20 个小时的历史数据, 抽取特征, 完成参数的学习, 来预测未来 5 个小时的 $PM2.5$ 浓度值。

模型输入为最近 20 个小时历史的测量数据 $x_1, x_2 \dots x_{20}$, 然后根据公式 (4) (5) (6) (7) (8) (9) 所述得到 $h_1, h_2 \dots h_{20}$ 输出隐状态向量, 经过公式 (10) 合并, 得到编码向量。

$$h = \text{Concat}(h_1, h_2 \dots h_{20}) \quad (10)$$

$$m = \sigma(W_{hm} h + b_{hm}) \quad (11)$$

$$PM2.5_o = W_{mo} m + b_{mo} \quad (12)$$

公式 (10) 中代表对向量首位相连拼接合并, 完成对历史时刻数据的编码。根据公式 (11) 对向量 h 经过一次全连接层线性变化得到模型最终的特征向量 m , 此时向量 m 就是模型对历史数据抽取的抽象特征。在公式 (11) 中 W_{hm} 和 b_{hm} 代表全连接层的权重和偏置。在输出层中, 如公式 (12) 所示, 模型会根据特征向量预测出未来时刻的 $PM2.5$, 用 $PM2.5_o$ 表示。 $PM2.5_o$ 是一个 5 维的实数向量, 包含预测出的未来 5 个小时的输出值 $PM2.5_{1h}, PM2.5_{2h}, PM2.5_{3h}, PM2.5_{4h}, PM2.5_{5h}$ 。

模型的训练阶段采用均方误差 MSE 作为的损失函数, 具体计算过程如公式 (13) 所示:

$$loss = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^5 (PM2.5_j - y_j)^2 \quad (13)$$

式中, N 代表样本的总数, j 代表未来第小时, $PM2.5_j$ 代表第 j 小时的预测值, y_j 代表第 j 小时的真实值。梯度优化过程中使用 Adagrad^[10] 策略进行参数更新, 初始的学习率设置为 0.01。

算法模型的完整训练过程如算法 1 所述。

算法 1: 基于 LSTM 的 $PM2.5$ 预测算法

- 1: 初始化 LSTM 神经网络参数, (公式 (4) (5) (6) (7) (8) (9) 中的 W 和 b)
- 2: 初始化全连接层网络参数, W_{hm} 和 b_{hm}
- 3: 初始化输出层网络参数 W_{mo} 和 b_{mo}
- 1: 定义 M 为训练集总共 batch 数
- 2: for each iteration $i=1, 2, \dots, M$ do
- 3: 采样一个 batch 的训练样本序列 x, y
- 4: 根据公式 (4) - (9) 得到历史序列隐状态序列 h_1, h_2, \dots, h_{20}
- 5: 根据公式 (10) 串联合并 h_1, h_2, \dots, h_{20}
- 6: 根据公式 (11) (12) 得到预测结果 $PM2.5_{1h}, PM2.5_{2h} \dots PM2.5_{5h}$
- 7: 根据公式 (13) 计算模型损失
- 8: 计算网络所有参数值的梯度
- 9: 更新参数值
- 10: end for

3 实验结果与分析

3.1 数据集

本文选择网络公开的北京空气质量数据作为实验评测数据, 数据包含 2014 年 1 月 1 日到 2014 年 12 月 31 日北京市内 35 个数据收集地点每小时的采集数据。实验按照 25 个小时大小的滑动窗口切分数据, 取前 20 个小时数据作为模

型输入数据，后 5 个小时的数据作为预测数据，生成样本对数据集。在样本对数据集中，本文随机选择 20% 作为测试数据，剩下的 80% 作为训练数据。舍弃含有空值和异常数据样本，共搜集得到 517 476 条训练样本，129 370 条测试样本。

3.2 实验结果

本实验环境主要参数 CPU: Intel (R) Core (TM) i7—6800K CPU @ 3.40GHz, GPU 显卡: 1 080 ti4, 内存为 32 GB, 操作系统为 Ubuntu 16.04 64 bit。实验设置训练迭代 epochs 为 100, LSTM 隐状态维度为 128 维, 损失函数选择回归预测常用的 MSE 均方误差, batch size 设置为 64。迭代完成 100 次训练, 模型训练学习曲线如图 4 所示。

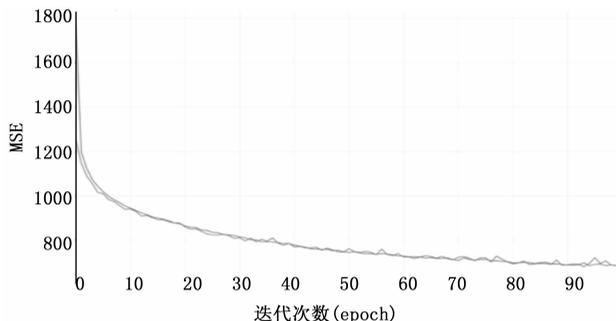


图 4 模型训练学习曲线

由图可知, 随着迭代次数的增加, 训练误差和测试误差都下降很快。当到迭代次数 (epoch) 达到 100 时, 测试误差和训练误差明显有分叉的迹象, 为避免模型发生过拟合, 本文选择到迭代次数到 100 次停止继续训练。可以看出, 测试集误差不断下降, 说明算法模型学到了数据的真实分布和潜在规律。

为验证所提出算法的有效性, 文本选取了 3 种机器学习算法作为对比参考, 分别是 ARIMA^[11], SVM^[12] 和 GBDT^[13]。对比实验各个算法根据给定大小的训练数据集训练模型, 在同样的测试集上得到测试结果, 以 MSE 和准确率做为评价标准。准确率以不同的相对误差作为条件, 给定最近 20 个小时的历史观测数据预测未来 5 个小时的 PM2.5 浓度。误差实验结果如表 1 所示。

表 1 MSE(均方误差)实验结果表

算法	MSE
ARIMA	899.309
SVM	997.574
GBDT	627.372
LSTM	525.537

在实验结果中 LSTM 取得了最小的均方误差, 在误差评价指标上表现出较好的预测能力。准确率实验结果如表 2 所示, 在实验过程中分别考虑了相对误差在 0.1, 0.2 和 0.3 的范围内, 对比不同算法取得的准确率, LSTM 同样取得了最好的预测效果, 在准确率上表现出良好的时序列预

测能力。

表 2 准确率实验结果表

算法	相对误差		
	0.1	0.2	0.3
ARIMA	0.379	0.610	0.770
SVM	0.337	0.606	0.731
GBDT	0.423	0.645	0.752
LSTM	0.460	0.668	0.797

为验证本算法模型稳定性和实用性, 本文选取测试数据中三个不同数据采样地点并随机采样 20 天数据作为时间轴对比不同的算法预测效果, 实验结果如图 5、图 6 和图 7 所示。从三个采样地点预测值与真实值对比来看, LSTM 预测值与真实值的拟合更加好, 没有出现较大的明显偏差。在图 5、图 6 和图 7 中, ARIMA, SVM 和 GBDT 三个对比预测算法均出现了在序列拐点处的滞后性, 与真实值有较大的偏差, 说明三个预测算法, 直接用历史数据作为特征或者对历史数据的简单处理作为特征, 并不能表现出其更深层次的序列规律。LSMT 也表现出神经网络特有的优势, 通过多种非线性运算能够对高维数据抽取抽象和深层的特征, 展现出比较好的预测和拟合效果。通多对同一种算法在不相同采样地点的横向对比, 发现本文提出的 LSTM 预测算法具有更好的稳定性。如对比 SVM 算法分别在图 5、图 6 和图 7 中的表现, 可以看到 SVM 在图 7 中的拟合效果要比图 5 和图 6 中的效果差很多, 而 LSTM 算法在三个图中均表现出了不错的效果, 没有根据不同采样地点数据的变化而受到很大的影响。

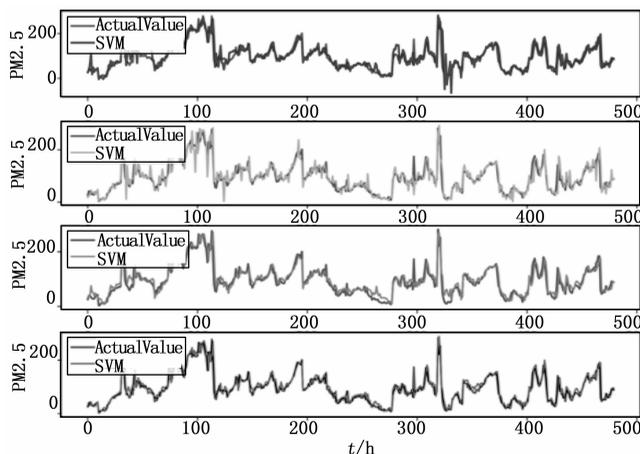


图 5 算法预测对比图 (前门观测点数据)

4 结束语

文本提出了基于深度学习 LSTM 循环神经网络的 PM2.5 预测算法, 算法模型不需要对数据进行复杂的专业处理, LSTM 可以抽取数据的时序特征, 对未来时刻 PM2.5 进行有效的回归预测。实验结果也表明, 本文提出的预测模型得到了良好的预测效果, 可以依据某个为位置的最近 20 小时历史测量数据, 对未来 5 个小时的 PM2.5 值

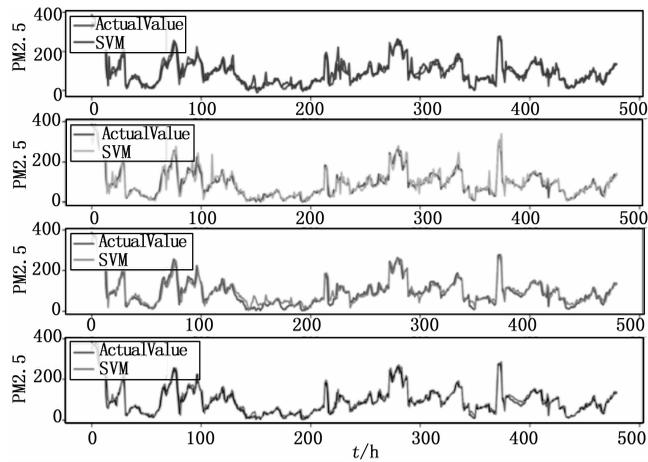


图 6 算法预测对比图 (天坛观测点数据)

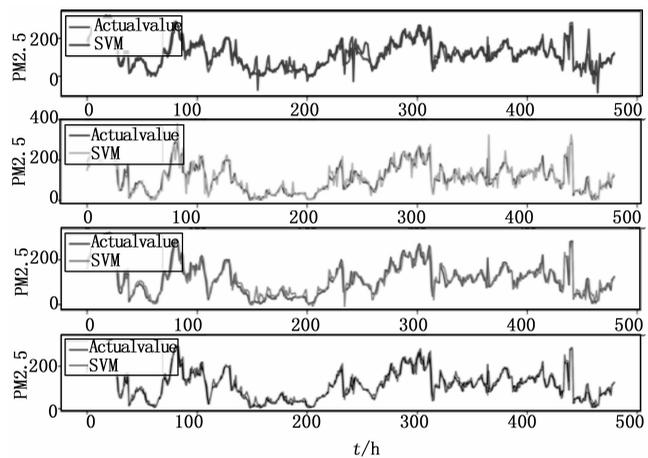


图 7 算法预测对比图 (通州观测点数据)

进行有效的预测。未来研究工作中可以对 PM_{2.5} 浓度变化规律和机理进行深入的分析研究, 结合更多的理论基础对 PM_{2.5} 进行更加精准的长时间段预测。

参考文献:

[1] LeCun Y, Bengio Y, Hinton G. Deep learning [J]. Nature, 2015, 521 (7553): 436 - 436.

(上接第 214 页)

分, 分别以 1 阶和 5 阶谐波为目标转换波, 验证了在最佳驱动条件下 DP-MZM 的波长转换方案能够实现最优的波长转换性能。以 16QAM-OFDM 信号为用户信号, 解调的电信号和发射电信号可以很好的归一化重合, 证明了基于 DP-MZM 的波长转换方案的传输稳定性和可行性。

参考文献:

[1] 胡文文. 发电厂通信组成及维护 [J]. 电气技术, 2011 (3): 48 - 50.
 [2] 严安. 水电厂通信光缆存在的问题及对策 [J]. 中国科技纵横, 2015 (2): 31 - 32.
 [3] 王德文, 刘杨. 一种电力云数据中心的任务调度策略 [J].

[2] Hochreiter S, Schmidhuber J. Long Short-Term Memory [J]. Neural Computation, 1997, 9 (8): 1935 - 1780.
 [3] Cobourn W G. An enhanced PM_{2.5} air quality forecast model based on nonlinear regression and back-trajectory concentrations [J]. Atmospheric Environment, 2010, 44 (25): 3015 - 3023.
 [4] Zhou Q, Jiang H, Wang J, et al. A hybrid model for PM_{2.5} forecasting based on ensemble empirical mode decomposition and a general regression neural network [J]. Science of the Total Environment, 2014, 496: 264 - 274.
 [5] 李嵩, 王冀, 张丹闯, 等. 大气 PM_{2.5} 污染指数预测优化模型仿真分析 [J]. 计算机仿真, 2015, 32 (12): 400 - 403.
 [6] Sahu S K, Mardia K V. A Bayesian kriged Kalman model for short-term forecasting of air pollution levels [J]. Journal of the Royal Statistical Society: Series C (Applied Statistics), 2005, 54 (1): 223 - 244.
 [7] Kurt A, Oktay A B. Forecasting air pollutant indicator levels with geographic models 3 days in advance using neural networks [J]. Expert Systems with Applications, 2010, 37 (12): 7986 - 7992.
 [8] Diaz-Robles L A, Ortega J C, Fu J S, et al. A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: The case of Temuco, Chile [J]. Atmospheric Environment, 2008, 42 (35): 8331 - 8340.
 [9] Chung J, Gulcehre C, Cho K H, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling [J]. arXiv preprint arXiv: 1412.3555, 2014.
 [10] Duchi J, Hazan E, Singer Y. Adaptive Subgradient Methods for Online Learning and Stochastic Optimization [J]. Journal of Machine Learning Research, 2011, 12 (7): 257 - 269.
 [11] Jenkins G M. Autoregressive - Integrated Moving Average (ARIMA) Models [J]. Encyclopedia of Statistical Sciences, 2004, 1.
 [12] Hearst M A, Dumais S T, Osuna E, et al. Support vector machines [J]. IEEE Intelligent Systems and their applications, 1998, 13 (4): 18 - 28.
 [13] Friedman J H. Stochastic gradient boosting [J]. Computational Statistics & Data Analysis, 2002, 38 (4): 367 - 378.
 电力系统自动化, 2014 (8): 61 - 66.
 [4] 王长乐. 射频电控光波长变换技术的研究 [D]. 天津: 天津大学, 2012.
 [5] 韩柳燕. 新型调制格式全光波长转换技术和全光逻辑门的研究 [D]. 北京: 清华大学, 2009.
 [6] Nuzman C, Leuthold J, Ryf R, et al. Design and implementation of wavelength-flexible network nodes [J]. Journal of Lightwave Technology, 2003, 21 (3): 648 - 663.
 [7] Yoo S J B. Wavelength conversion technologies for WDM network applications [J]. Journal of Lightwave Technology, 1996, 14 (6): 955 - 966.
 [8] Zhu D, Yao J. Dual-Chirp Microwave Waveform Generation Using a Dual-Parallel Mach-Zehnder Modulator [J]. IEEE Photonics Technology Letters, 2015, 27 (13): 1410 - 1413.