

Traffic speed prediction using deep learning method

文章结构

文章贡献

Introduction

文章动机

模型架构

数据集

数据预处理

实验与讨论

模型参数

测试结果

文章结构

- Introduction
 - 交通流速预测在交通研究领域的定位
 - 交通流信息预测方法的研究流派及各流派已有研究成果
 - 作者写这篇文章的动机
 - 文章框架
- Methodology
 - DBN原理与介绍
 - 用于交通流速预测DBN网络结构
 - 性能指标和预测时段长度选择
- Data Description
 - 数据来源
 - 数据结构
 - 训练集/测试集划分
 - 数据预处理（不同车道流速怎么聚合）
- Experiment and Discussion
 - 模型超参数：**网格搜索**，以MAPE为指标利用训练集确定最优参数
 - 输入序列长度

- 隐藏层层数（RBN+BP层）
- 隐藏层各层神经元数
- epoch数（迭代次数）
- 实验结果
 - 不同模型性能对比：DBN优于ARIMA和BPNN
 - DBN在不同预测长度下效果：预测长度越长预测效果越差，尤其体现在流量高峰期，偏差较大
- Summary

文章贡献

提出使用深度信念网络（DBN）来进行短期交通流速度预测，对比了以自回归滑动平均模型（ARIMA）为代表的**传统含参数统计模型**和以反向传播网络（BPNN）为代表的**无参数浅层机器学习模型**在不同预测时间跨度下的预测效果。实验结果证明了DBN在交通流信息预测上是很有效的，改进了传统方法不能反映交通流随机性特征的缺陷。

此外，给出了DBN模型在时空网络上的一种较优的超参数组合。

Q1：哪里体现出传统方法在反映交通流随机性特征上有困难？

文章没有提这种随机性特征包含什么，只能体现在预测效果上（论证不严谨）

Q2：为什么说机器学习模型以NN为代表是无参数的？这里的参数是指什么？

- 非参数模型≠模型中没有参数，而是参数很多/参数不确定。

non-parametric类似单词priceless，并不是没有价值，而是价值非常高、无价，也就是参数非常多（注意：所谓“多”的标准，就是参数数目大体和样本规模差不多）。通过有限个参数来确定一个模型，就称为“有参数模型”，如线性回归、Logistic回归。

- 参数模型：明确指定了学习到的目标函数形式，即作了形式假设

比如线性回归模型（一次方程），通过训练学习到具体参数。假设可以简化学习过程，但也会限制学习的内容，简化目标函数为已知形式的算法就称为**参数机器学习算法**。通过固定大小的参数集(与训练样本数独立)概况数据的学习模型称为参数模型。

Introduction

交通流速预测是交通信息预测的子领域。作者将当前交通信息的预测方法分为两个主要流派：参数模型（parametric model）和非参数模型（non-parametric），并分别介绍了不同流派的代表模型以及已有研究取得的成果。

作者着重指出，随着近年来深层神经网络的兴起，深度学习方法逐渐被用于交通信息的预测中，并被证明在大规模数据集上能够取得很好的效果。特别地，由于深度信念网络（DBN）能够在无监督情况下很好地提取数据中的非线性关系，已经开始被应用于交通流预测中。

参数模型主要包含：

- 自回归滑动平均模型（ARIMA）及其衍生模型（ARIMAX/SARIMA）
- 卡尔曼滤波（Kalman Filter）为代表的状态空间模型（state-space models）。

非参数模型主要包含一些机器学习模型，特别是浅层神经网络：

- 反向传播网络（BPNN）
- 长短时记忆神经网络（LSTM NN）

这些模型相比于传统的参数模型，能够预测的时间范围更长，效果也相对更好。

文中提到的深度学习模型则包括：

- 栈式自编码器（Stacked Autoencoder, SAE）
- 深度信念网络（Deep Belief Network, DBN）

已有研究在DBN上使用多任务回归层，在交通流预测上取得不错效果。

BPNN和LSTM NN都是只有一层的神经网络，其中LSTM NN可以进行较长期预测，解决了BPNN的梯度衰减问题。

LSTM NN is composed of one input layer, one LSTM layer with memory blocks, and one output layer. Because LSTM NN can automatically calculate the optimal time lags, and thus no predetermined time window size is needed.

Deep Learning 方面：对于大规模数据集很管用

- SAE（Stacked Auto Model）：预测交通流与识别每辆车的车速
- DBN（deep belief network, 深度置信网络）：已有研究使用多任务回归层将其用于交通流预测取得不错成果

文章动机

现有研究没有对比过DBN与传统统计参数模型以及浅层机器学习方法在交通流速预测上的具体效果。

模型架构

直接采用DBN网络。输入 \mathbf{x} 为预测时间节点前的交通流速序列向量，输出 \mathbf{y} 为预测范围内的交通流速序列向量。

关于RBN与DBN的笔记：

[RBM与DBN.pdf](#)

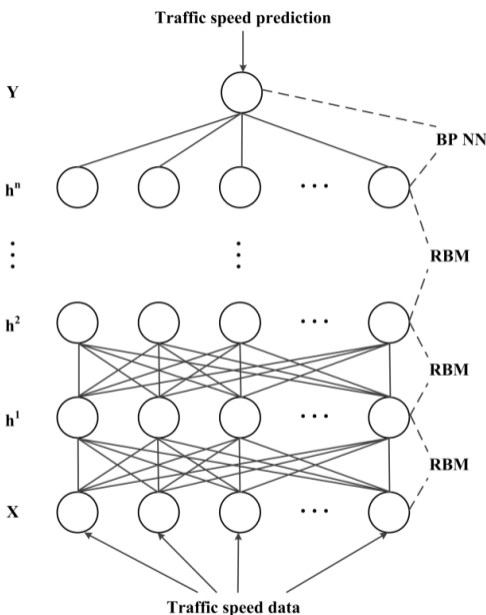


Figure 1. Structure of the DBN model.

训练过程：

- 预训练：使用无监督学习方法训练每一层RBN（受限玻尔兹曼机）
- 微调：再通过部分带标签数据，使用反向传播算法整体优化模型

数据集

选用了2013年6月至2013年8月期间的北京二环三环间主干道（德胜门–马甸桥）交通流数据，通过探测器分别采集各车道数据，时间间隔为2min一次，包含速度、流量、占有率三大特征。

划分数据集如下：

- 训练集：6月1日–8月24日
- 测试集：8月25日–8月31日

数据预处理

对各车道用加权方式进行聚合，得到输入输出的标准交通流速：

$$\text{Speed}_{\text{segment}} = \frac{\sum_i (\text{Flow}_i \times \text{Speed}_i)}{\sum_i \text{Flow}_i}$$

其中， $\text{Flow}_i, \text{Speed}_i$ 分别代表第*i*个车道上的交通流量与速度。

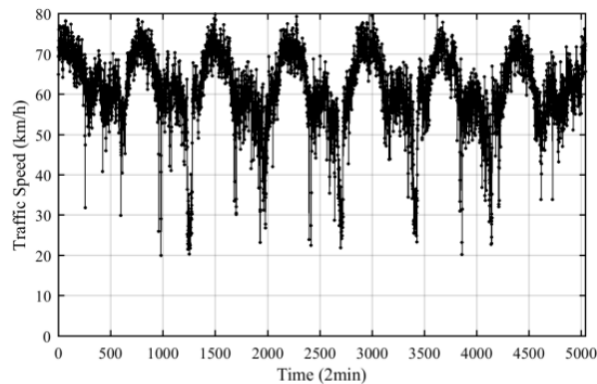


Figure 2. Traffic speed data on study segment for some week.

实验与讨论

模型参数

模型的超参数主要包含：

- 输入序列长度（previous intervals）
- 隐藏层层数（layer number）
- 隐藏层各层神经元数（layer units）
- 训练迭代次数（epochs）

采用网格搜索方法，以MAPE（平均绝对百分比误差）为评价指标在训练集上进行测试，得到不同预测时间范围下的最优超参数组合：

TABLE I. OPTIMAL PARAMETERS FOR THE DBN MODEL.

<i>Prediction horizon</i>	<i>Previous intervals</i>	<i>Layer number</i>	<i>Layer units</i>	<i>Epochs</i>
2-minute	8	1	500	60
10-minute	12	1	400	85
30-minute	25	1	200	45

参数组合结果表明，对于当前较小的北京干线数据集，多个RBN层并不能取得更好的效果，一层RBN已经足够对其时空特征进行建模。但对于更大的数据集，或许需要使用更复杂的网络架构。

测试结果

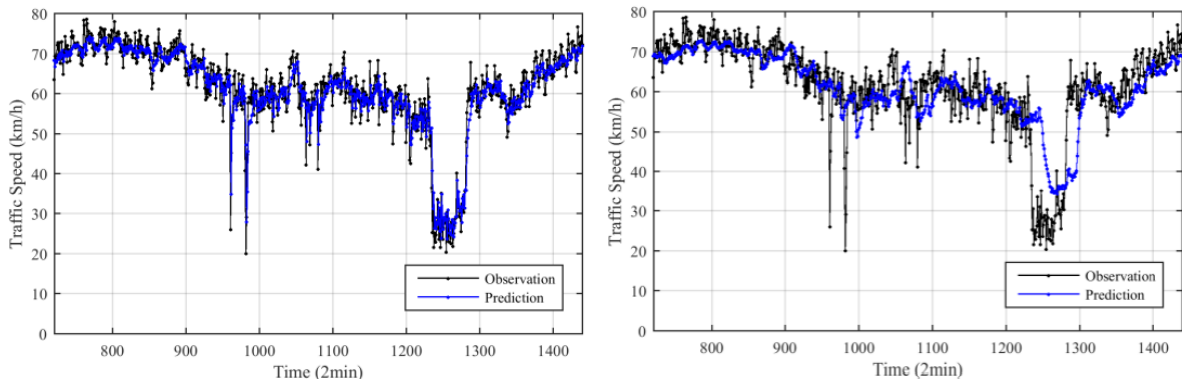
分别测试了2min、10min和30min下DBN、ARIMA与BPNN的预测效果。实验结果证明：在不同预测时间跨度下，DBN的预测效果均优于ARIMA和BPNN（分别代表传统参数统计方法以及浅层机器学习方法）。

TABLE II. EXPERIMENT RESULTS

<i>Prediction horizon</i>	<i>Model</i>	<i>MAPE</i>	<i>RMSE</i>	<i>RMSN</i>
2-minute	DBN	5.8099%	4.3585	0.0710
	BPNN	5.9681%	4.4880	0.0731
	ARIMA	5.9752%	4.5116	0.0735
10-minute	DBN	7.3359%	5.5365	0.0902
	BPNN	7.5277%	5.7688	0.0940
	ARIMA	7.8387%	6.0834	0.0991
30-minute	DBN	8.4782%	6.3109	0.1029
	ARIMA	8.8433%	6.5561	0.1069
	BPNN	8.9312%	6.9797	0.1137

其他实验结论：

随着预测时间长度的增加，三种预测方法预测结果的准确度都在下降，即：更多交通流的随机特征被丢失。这种准确度的下降尤其体现在早晚高峰时段。



（左右图分别为一天中预测时间长度分别为2min和30min的预测值和观测值对比）

