

基于改进双层 LSTM 的轨道交通客流预测分析研究

刘云 万心宇

(武汉地铁运营有限公司 湖北武汉 430030)

摘要:近十多年来,我国轨道交通行业飞速发展,各大一二线城市均成立了当地的地铁公司,助力各城市的轨道交通实现了从无到有、从小到大的迅猛发展。随着轨道交通行业日益壮大,客流预测逐渐成为各地铁公司必要的工作内容之一,准确实时的客流预测不仅能让地铁公司及时准确地对列车运行与人员分配做出合理安排,更能在客流高峰期来临前及时预警,避免发生生产安全事故。轨道交通客流数据是一种典型的低维时间序列数据,针对此类数据,本文提出将深度学习中的长短时记忆网络 LSTM 应用于客流预测,同时根据现实场景与其他相关客流影响因素,引入机器学习模型对结果进行纠偏,从而实现准确及时的客流预测,辅助地铁公司进行相关工作安排与决策。

关键词:轨道交通客流预测;时间序列;长短时记忆网络;机器学习

轨道交通行业在近年来取得了长足的发展与进步,一条条地铁线路成为城市的主动脉,不仅为城市居民的出行提供了便利,也带动了地铁沿线经济的发展,成为城市经济发展中不可缺少的一分子。仅以武汉为例,武汉地铁自 2012 年 2 号线开通运营开始到如今,地铁线路已然增至 11 条,车站数量也从十年前的 26 座增加至 282 座。轨道交通行业作为我国的民生行业,其运营稳定性关乎城市中数百万人的正常出行,对我国的国计民生有着重要的影响。客流预测作为其中的重要内容,关系着各地铁公司相关工作计划的安排与正常工作的开展。

一、相关工作

客流预测的方法有很多种,目前国内外的研究大致有几类,如以回归模型为代表的自回归综合移动平均模型 (Autoregressive Integrated Moving Average model, ARIMA)、非线性回归模型 (Nonlinear Auto-Regressive, NAR);以分类算法模型为代表的全连接神经网络模型。还有如四阶段预测法,具体为将预测分为四个阶段:交通生成预测、交通分布预测、交通方式划分预测和交通分配,从而生成最终的预测客流。各类方法都可以实现客流预测,但在使用场景上存在局限性。总体来说,通过时间序列预测的方式实现客流预测是当前的主流方法。近年来,随着深度学习相关理论的发展,许多深度学习模型被应用到了客流预测中,也实现了良好的效果。

本文主要研究地铁客流数据,考虑到客流数据是一种典型的低维时间序列数据,其数据具有明显的趋势性与周期性,因此研究将深度学习中的长短时记忆网络 (Long Short-Term Memory Network, LSTM) 模型应用于客流预测。同时,考虑到地铁客流也会受到其他偶发因素的干扰,比如政策因素 (根据政策需求关停相关站点)、气候因素 (天气恶劣时出行人流较少) 等,本文研究在 LSTM 模型进行客流预测的基础上,引入其他会对地铁客流产生较大的影响的偶发因素,建立机器学习模型,对预测结果进行纠偏,从而提升客流预测的准确度。

二、深度学习与长短时记忆网络

1. 深度学习

早些年,在数据分析挖掘领域,专家系统与传统机器学习算法占据行业主流,应用于各个工业领域,展现了良好的效果。然而近十几年来,由于数据量的指数式增长,数据呈现出了大规模、多样性、低价值密度等特点,传统数据挖掘方法难以适应现实需求。同时,随着计算机算力的增长,深度学习逐渐兴起,成为了人工智能的主要研究方向,其在工业界成功应用的案例也比比皆是。比如:以卷积神经网络 (Convolutional Neural Networks, CNN) 为代表的神经网络模型,在人脸识别、目标检测、图像分类等领域获得了广泛的应用;以循环神经网络 (Recurrent Neural Networks, RNNs) 为代表的模型,则在语音识别、文本翻译、情感分类等时序数据领域中取得了成功。在 RNN 模型中, LSTM 又是其中最具有代表性的网络结构之一。

本文主要研究地铁客流数据,对未来客流进行实时预测报警。考虑到时序数据的特点,本文研究将 RNNs 网络中的 LSTM 网络应用于客流预测,借助于 LSTM 网络能处理长时间的时序数据的能力,进行数据预测。

2. 循环神经网络

循环神经网络是一种典型的处理序列数据的深度神经网络。比起传统的人工神经网络 (Artificial Neural Network, ANN), 它的特点在于加入了“记忆”的成分。如 (x_1, x_2, \dots, x_m) 可以被视作一个时间序列数据 (或者说随着时间不断发展的事件序列), 如果用传统的全连接人工神经网络进行数据处理, 此网络会对这些输入分别进行计算, 得到相应的 (o_1, o_2, \dots, o_m) 一系列的输出, 但是 o_1, o_2, \dots, o_m 之间并无关联。循环神经网络的特点在于: o_2 的值既与 x_2 相关, 也与 o_1 的值有关。循环神经网络的网络结构如图 1 所示。

图 1 循环神经网络结构图

Fig.1 The structure of RNNs

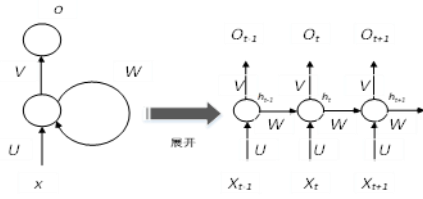


图1的左图为循环神经网络的折叠形式，右图为展开形式。在此结构中， x 表示输入的序列数据， h 表示网络的隐藏层状态， o 表示神经元的输出向量， U 为输入层到隐藏层之间的参数矩阵， V 为隐藏层到输出层的参数矩阵， W 为不同时刻隐藏层之间的参数矩阵，所有参数矩阵均为不同时间隐藏状态的共享矩阵。该网络的前向传播过程可以公式化表示为：

$$a_t = b + Wh_{t-1} + Ux_t \quad (1)$$

$$h_t = \tan(a_t) \quad (2)$$

$$o_t = c + Vh_t \quad (3)$$

$$\hat{y}_t = \text{soft max}(o_t) \quad (4)$$

在上述公式(1) - (4)中， x_t 为输入数据在 t 时刻的输入， h_t 为隐藏层 t 时刻的状态表示， a_t 为隐藏层在 t 时刻的输入， o_t 为 t 时刻的输出， \hat{y}_t 为预测值经过归一化之后的概率输出。

3. 长短时记忆网络

尽管普通的循环神经网络已经能做到对序列数据的有效处理，但由于梯度消失与梯度爆炸的原因，导致其只能处理短时间的序列，无法对长段时间的序列进行有效处理。正是为了解决这个问题，LSTM网络出现了。相比一般的RNN网络，LSTM的主要改进在于多了遗忘门、输入门、输出门，这三个控制门的设计。其网络结构如图2所示。

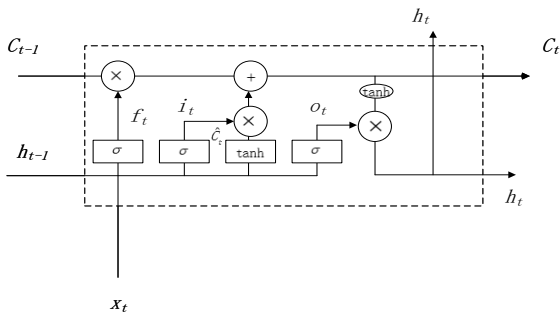


图2 LSTM 细胞结构

Fig.2 The cell structure of LSTM

此结构为LSTM网络单个时刻的细胞结构图，类似于普通RNN网络，LSTM同样也是以此细胞结构为单元展开。其中， x_t 表示当前时刻的输入，其前向传播的公式化表示为：

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \quad (5)$$

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (6)$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (7)$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \quad (8)$$

其中，在公式(6)中， f_t 表示遗忘门，它的输入为上一时刻的隐层状态 h_{t-1} 与当前时刻的输入 x_t ，取值范围在0到1之间，控制着上一时刻信息 C_t 的通过程度。公式(7)中的 i_t 表示输入门，控制着新记忆 \tilde{C}_t 写入长期记忆的程度。LSTM网络中这种跨层连接的设计有效地减弱了梯度消失的影响，使其对于长序列的记忆预测能力大大提高。正是由于LSTM，使得我们可以将其应用于地铁客流的预测模型中。地铁客流数据具备明显的周期性与趋势性，为保证预测模型的准确性，其模型的输入数据需要至少长达一年的数据，以保证其输入数据的长度至少包含一个完整的周期，这正是典型的长片的时间序列预测问题。

三、LSTM 在地铁客流预测中的应用

近些年来，RNN类型网络在文本分类、情感分析、机器翻译等序列数据处理领域取得了极大的成功，因而也开始有人将其运用于地铁客流的预测问题。此处，本文首先介绍如何将LSTM模型应用于地铁客流预测，以及如何选择合适的LSTM模型。然后，本文会继续分析如何基于预测数据，做进一步的数据纠偏，得到更为准确的客流预测数据。

1. 将双层LSTM应用于客流预测

地铁客流数据是一个典型的单维时间序列数据，在建立模型之前，需要完成数据的标准化与准备工作。将客流数据符号化表示为 $X = (X_1, X_2, \dots, X_n)$ ，其中 X_t 表示时序数据 X 在 t 时刻的值，其为一个 m 维的向量，其符号化表示公式为：

$$X_t = \{x_t^{(1)}, x_t^{(2)}, \dots, x_t^{(m)}\}, X_t \in R^m \quad (9)$$

通过将历年客流数据汇总，以一年为时间片长度，将客流数据片段为一段一段的时间序列数据片段，然后进行填充空值处理、标准化处理，将数据整理为规范化的待使用数据集。在训练时，考虑到单层LSTM模型的模型复杂度不足以满足现实地铁客流数据复杂性的需要，因为采用双层LSTM模型进行模型训练，这样既是为了允许更大的模型复杂性、增大模型容量，也是为了防止过拟合，确保模型与现实场景的地铁客流数据相匹配。在训练过程中，以双层LSTM模型为基础，根据可调参数，将模型训练分为不同的训练组，并划分训练集、测试集与验证集。通过训练模型，对比准确率，从而得到合适的特定参数的LSTM模型。此处，关于模型准确率的评价，首先引入以下四个参数：

TP ：将样本判定为故障样本，且实际为故障样本的数量；

TN ：将样本判定为正常样本，且实际为正常样本的数量；

FP ：将样本判定为故障样本，且实际为正常样本的数量；

FN ：将样本判定为正常样本，且实际为故障样本的数量；

相应的评价指标计算方式如下。

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (10)$$

$$precision = \frac{TP}{TP + FP} \quad (11)$$

$$recall = \frac{TP}{TP + FN} \quad (12)$$

2. 基于客流预测结果的机器学习模型

在客流预测中,单纯依赖序列数据的预测模型进行客流预测是一个过于理想的情况,因为在训练模型之时,本身就只考虑了客流数据本身以及前后不同时间片的客流数据之间的关联性,并没有考虑客流数据之外的其他因素,最典型的有比如政策影响、天气影响、安全事故影响等。因此,单纯依赖 LSTM 实现客流预测,必须在一个绝对理想的数学环境下才能实现高精度与高准确率,一旦现实情况出现偏差,预测结果也将迅速偏离实际值。

在实际模型建立的过程中,考虑根据实际工作经验,将客流数据之外的其他会影响客流的数据纳入到客流预测模型中,比如 0-1 化之后的节假日数据、气象数据等,分析此类数据的影响程度,选择数个影响力最大的数据,进行标准化和去冗余操作,形成完备的训练数据集。然后选择合适的分类聚类模型,比如支持向量机 (Support Vector Machine, SVM)、K 均值 (K-means) 聚类算法、BP (Back Propagation) 网络等。比如以筛选出的其他指标数据为输入,以预测误差为输出,训练分类模型,在进行客流实时预测时,在 LSTM 预测客流的基础上,实时以当日或次日的其他指标数据为输入,计算出预测误差,从而对预测结果进行实时的纠偏,得到更为准确的客流预测结果。

四、结语

本文主要研究将深度学习中的 LSTM 网络应用于地铁客流预测中,提出了一种基于改进双层 LSTM 模型的轨道交通客流预测思路与方法。

该模型适用于轨道交通客流这类典型的单维度长时间片时序数据,既能学习历史数据,根据过往数据预测未来客流量,也能做到考虑各种可能影响客流的因素,进行预测结果的实时纠偏。既填补了轨道交通领域的相关理论,也顺应了轨道交通行业能化发展趋势,具有较强的现实意义。

参考文献:

- [1]Liu Y,Liu Z,Jia R.DeepPF:A deep learning based architecture for metro passenger flow prediction[J].Transportation Research Part C:Emerging Technologies,2019,101:18-34.
- [2]Zhang Z,Wang C,Gao Y,et al.Short-Term Passenger Flow Forecast of Rail Transit Station Based on MIC Feature Selection and ST-LightGBM considering Transfer Passenger Flow[J].Scientific Programming,2020:1-15.
- [3]Wei Y,Chen M C.Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks[J].Transportation Research PartC:Emerging Technologies,2012,21(1):148-162.
- [4]Liu S,Yao E.Holiday passenger flow forecasting based on the modified least-square support vector machine for the metro system[J].Journal of Transportation Engineering,PartA:Systems,2017,143(2):04016005.
- [5]Zhang Z,Wang C,Gao Y,et al.Short-term passenger flow forecast of rail transit station based on MIC feature selection and ST-LightGBM considering transfer passenger flow[J].Scientific Programming,2020