

# 基于TCN和LightGBM的供应商主动维护客户构成模型

刘 铭, 何利力, 郑军红\*

浙江理工大学计算机科学与技术学院, 浙江 杭州

收稿日期: 2023年4月20日; 录用日期: 2023年6月27日; 发布日期: 2023年7月5日

## 摘 要

本文介绍了一种基于TCN和LightGBM的供应商主动维护客户构成模型,旨在解决传统的被动响应式服务无法满足当今竞争激烈市场需求的问题。该模型结合了时间序列分析和机器学习技术,使用TCN捕捉客户订单数据的时间依赖性,同时利用LightGBM学习非时序的复杂关系特征,以快速捕捉市场变化并提高服务体验,让客户既不断货,也不积压,实现“响应性服务”向“感知响应性主动服务”转变。本文的贡献在于提出了一种高效的供应商维护客户构成模型,具有较强的适应性和预测准确性。该模型的应用可以有效地提高供应商的竞争力,满足市场的需求。

## 关键词

TCN, LightGBM, 主动服务, 供应链

# A Customer Composition Model for Supplier Proactive Maintenance Based on TCN and LightGBM

Ming Liu, Lili He, Junhong Zheng\*

School of Computer Science and Technology, Zhejiang Sci-Tech University, Hangzhou Zhejiang

Received: Apr. 20<sup>th</sup>, 2023; accepted: Jun. 27<sup>th</sup>, 2023; published: Jul. 5<sup>th</sup>, 2023

## Abstract

This article introduces a customer composition model for supplier proactive maintenance based

\*通讯作者。

on TCN and LightGBM, aiming to solve the problem that traditional passive responsive services cannot meet the demands of today's competitive market. The model combines time series analysis and machine learning techniques, using TCN to capture the time dependency of customer order data and utilizing LightGBM to learn non-time series complex relationship features, in order to quickly capture market changes and improve service experience, allowing customers to have continuous supply without inventory pile-up, achieving the transformation from "responsive service" to "perceptive responsive proactive service". The contribution of this article is to propose an efficient supplier maintenance customer composition model, which has strong adaptability and prediction accuracy. The application of this model can effectively improve the competitiveness of suppliers and meet market demands.

## Keywords

TCN, LightGBM, Active Service, Supply Chain

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着经济全球化和科技物流的迅速发展,现代企业供应链的科学有效管理已经离不开信息技术的支持。传统的按订单组织货源的供应模式存在一些问题,如库存过高、需求预测不准确、协调成本高、反应速度慢等,这些问题都严重影响了供应链的运作效率和企业的盈利能力。为了解决这些问题,企业需要主动维护客户模型,对客户需求进行精准预测,从而帮助供应商作出正确决策,提高供应链的运作效率和企业的盈利能力。

机器学习算法可以更好地处理非线性、非平稳、异方差等问题,并且具有更高的灵活性和准确性[1]。常用的模型有随机森林(Random Forest, RF) [2],支持向量回归(Support Vector Regression, SVR) [3],梯度提升树(Gradient Boosting Decision Tree, GBDT) [4]等方式。Candelieri 等[5]采用并行优化模型对 SVM 回归的超参数进行了优化,并在用水需求量数据集上进行了短期预测。Chen 等[6]提出了一种基于 XGBoost 的学习模型,其采用了多种优化技术,如特征列排序、缺失值处理、块压缩、直方图近似等,来加速模型的训练和预测,并且具有较好的泛化能力和鲁棒性。该模型被数据科学家广泛使用,并且具有高可扩展性,在众多领域取得了很好的成果。但是, XGBoost 仍然存在一些局限性,例如在处理大规模数据集时,模型的训练和预测速度可能会受到限制。为了解决这个问题,Ke 等[7]提出了 LightGBM (Light Gradient Boosting Machine)机器学习算法,主要是采用了一些与 XGBoost 不同的优化技术,如直方图采样、基于梯度单边采样、特征并行和数据并行等,来提高模型的训练和预测速度,并且具有较高的精度和泛化能力。王华勇等[8]将 LightGBM 算法应用于短期负荷预测,证明了该算法具有更高的计算效率以及计算精度。机器学习技术可以帮助企业提高销售预测的准确性和效率,从而更好地制定销售战略和决策[9]。

随着计算机计算能力的不断提高,深度学习模型能够自动地从大量的历史销售数据中学习到潜在的模式和趋势,并能够对未来的销售进行时序预测[10]。基于此, Hochreiter 等[11] 1997 年提出了 RNN 的改进模型——长短时记忆(Long Short-Term Memory, LSTM)网络模型。LSTM 是一种循环神经网络,常用于处理序列数据,如自然语言处理、语音识别、动作识别以及时序预测等任务。尽管 LSTM 在处理短序列方面表现出色,但在处理高维数据时可能会面临参数量过大的问题[12]。为了解决 LSTM 长时间记忆

和训练速度慢等问题,一种特殊的一维卷积神经网络时间卷积网络(Temporal Convolutional Network, TCN) [13]被提出并广泛应用于时序数据的处理中,它使用了一系列的卷积层和池化层,这使得它能够并行地处理时序数据,并且具有较少的参数和较高的训练效率。翟等[14]基于 TCN 算法对时序数据进行了分析和预测,并与其他相关模型在公共数据集上的验证进行了比较。结果表明,相比于 LSTM 模型,TCN 算法具有更好的预测效果。

本文利用数据挖掘技术挖掘各种影响订单量的相关特征,提出基于 TCN 和 LightGBM 的供应商主动维护客户构成模型,该模型使用 TCN 捕捉客户订单数据的时间依赖性,同时利用 LightGBM 学习非时序的复杂关系特征,以快速捕捉市场变化并提高服务体验,让客户既不断货,也不积压,实现“响应性服务”向“感知响应性主动服务”转变。

## 2. 基于 TCN 和 LightGBM 的供应商主动维护客户构成模型

### 2.1. 问题定义

$N$  条时间上相关的多变量序列可以表示为:  $Y = \{y_1, y_2, \dots, y_T\}^N$ ,  $\in \mathbb{R}^{N \times T}$ , 以  $y_{i,j} \in \mathbb{R}$  来表示在第  $i$  条时间序列  $j$  时刻的值。我们做多变量序列预测目标是在给定 TGE 时间步的序列值  $Y = \{y_1, y_2, \dots, y_T\}^N$  和模型参数  $\Theta$  的情况下预测出未来  $k$  个时间步序列的每个时刻的值,具体如式(1)所示:

$$\hat{y}_{T+tT+k} = f(y_{1:T} : \Theta) \tag{1}$$

其中,  $f(\cdot)$  为待建的预测模型,  $\Theta$  为模型的参数。

### 2.2. 时间卷积神经网络(TCN)

时序卷积网络(TCN)以卷积神经网络(CNN)模型为基础,使用一维卷积层对序列数据进行局部感知,然后通过堆叠多个卷积层来构建更复杂的模型。在每个卷积层中,卷积核的大小和步幅可以根据数据的特征进行调整。此外,TCN 还可以使用残差连接和扩张卷积等技术来进一步提高模型的性能。

#### 2.2.1. 因果卷积

因果卷积(Causal Convolution)是一种卷积操作,它可以在卷积核的前面添加一个因果滤波器(Causal Filter)来保证卷积操作是因果的,即输出只依赖于输入中的过去值,而不依赖于未来值。因果卷积的结构如图 1 所示。

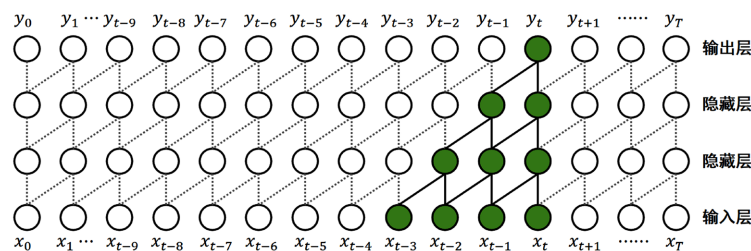


Figure 1. The structure of the causal convolution  
图 1. 因果卷积结构

因果卷积的输入是定义为  $X = \{x_0, x_1, \dots, x_t, \dots, x_T\}$ , 输出定义为  $Y = \{y_0, y_1, \dots, y_t, \dots, y_T\}$ , 卷积核为  $H = \{h_0, h_1, \dots, h_t, \dots, h_T\}$ , 如果使用普通的卷积操作,输出序列的每个值  $y_t$ , 具体如式(2)所示:

$$y_t = \sum_{k=1}^K h_k x_{t-k} \tag{2}$$

其中,  $K$  是卷积核的宽度,  $t$  是当前时刻。

但是, 这种卷积操作会使用当前时刻之后的输入值, 从而导致输出不是因果的, 因此不适用于序列数据的建模。为了解决这个问题, 可以使用因果卷积。

在因果卷积中, 卷积核的每个元素都只取决于输入序列中的过去值, 而与未来值无关。因此, 可以修改卷积操作的公式, 使用一个因果滤波器  $c$  来对输入序列进行调整, 具体如式(3)所示:

$$y_t = \sum_{k=1}^K h_k c_{t-k} \quad (3)$$

其中,  $c_t = x_t$ , 当  $t < K$  时,  $c_t$  表示在输入序列前面填充了  $K-1$  个 0 后的结果。这样,  $y_t$  就只依赖于  $c$  序列中的过去值, 而与未来值无关, 从而保证了输出序列的因果性。

由图 1 可知, 在因果卷积中, 输出层中的每个样本值都是由上一层中的一些先前样本值组成的加权和。这种加权和的权重随着时间的推移而减小, 因为输出信号的每个样本值只能依赖于先前的输入信号样本值。因此, 因果卷积中的权重只考虑过去的输入值, 而不会考虑未来的值。然而这样网络结构同样也会导致因果卷积对历史信息的覆盖范围不大。因果卷积的感受野随着网络的深度线性增加, 扩大模型的感觉, 需要堆叠较多的卷积层, 容易出现梯度爆炸和过拟合问题。

### 2.2.2. 扩张卷积

为了扩大因果卷积的扩大网络对输入数据的理解范围, 因此引入了扩张卷积。经过扩张后的因果神经网络结构如图 2 所示。

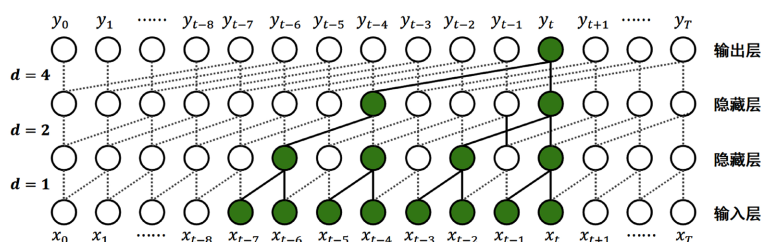


Figure 2. The structure of extended causal convolution

图 2. 扩张因果卷积结构

由图 2 可知, 在通过扩张卷积后的因果卷积神经网络, 下层神经元对于上一层神经元的历史信息的感觉野有明显的提升。对于一维数据的输入  $X = \{x_0, x_1, \dots, x_t, \dots, x_T\}$  和过滤器  $f = \{0, 1, 2, \dots, n-1\}$ , 时序  $T$  的扩张卷积操作  $H(*)$ , 具体如式(4)所示:

$$H(T) = (X *_d f)(T) = \sum_{i=0}^{n-1} f(i) \cdot x_{T-di} \quad (4)$$

其中,  $n$  表示卷积核大小,  $d$  表示扩张系数,  $T-d \cdot i$  表示对历史某一信息的定位。

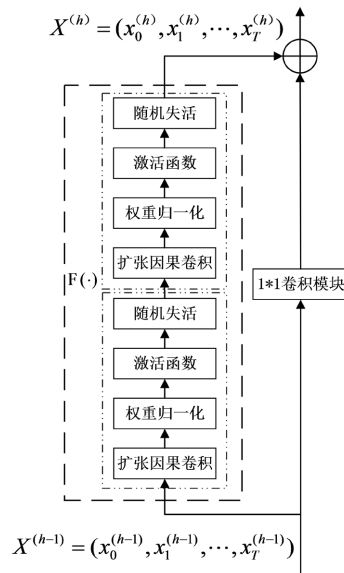
在不改变因果卷积神经网络的深度的情况下, 我们可以通过增加卷积核大小和扩张系数两种方法来增加了 TCN 的接受域。

### 2.2.3. 残差块

除了改变卷积核大小和扩张系数可以增加 TCN 的接收域, 还可以通过增加隐藏层的数量。然后, 深层次的网络会影响模型的训练和发生梯度消失。为了解决这一问题, TCN 引用了残差块来解决梯度不稳定的问题。残差块的结构如图 3 所示。

TCN 的残差块主要由两个分支模块组成, 一个分支主要包括一系列的卷积操作, 另一个分支执行一个简单的  $1 \times 1$  卷积, 来保持其输入通道的数量与扩张因果卷积的输出通道数量的一致性, 具体如式(5)

所示:

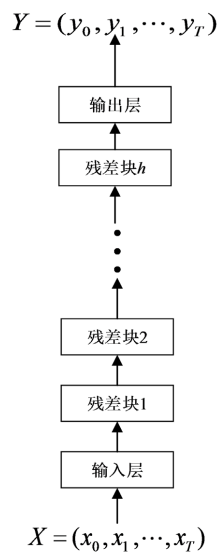


**Figure 3.** Internal structure of residual block  
**图 3.** 残差块内部结构

$$X^h = \delta\left(F\left(X^{h-1}\right) + X^{h-1}\right) \tag{5}$$

其中,  $\delta(\cdot)$  代表一个激活操作,  $F(\cdot)$  表示一系列的卷积操作, 主要包括双层的扩张卷积层、权重归一化 (Weight Normalization)、ReLU 激活函数和随机失活(Dropout)。

具体来说, 扩张卷积层主要由上述的因果卷积和扩张卷积组成, 用于提取潜在的特征。权重归一化主要用于归一化神经网络中的权重, 从而加速网络的收敛速度和提高泛化性能。ReLU 激活函数有非常好的收敛性和计算效率。随机失活可以减少过拟合的风险, 同时提高模型的泛化能力。



**Figure 4.** Deep TCN model  
**图 4.** 深层次的 TCN 模型

图 4 所示的是一个由  $h$  个残差模块叠加形成的深层 TCN。相比于浅层网络，深层网络具有更强的表达能力，因为它可以逐渐提取复杂的特征通过多个非线性层。因此，将多个残差模块堆叠起来可以构建一个更深的网络，从而提高模型的表达能力，使得模型能够更好地拟合时间序列数据。这种深层 TCN 模型具有更好地捕捉时间序列中的长期依赖关系的能力。

### 2.3. LightGBM 模型

LightGBM 是一个基于决策树和提升思想的梯度学习框架，由 Ke 等人于 2017 年提出，被广泛用于分类和回归任务。在本文中，我们采用了单边梯度采样(Gradient-Based One-Side Sampling, Goss)排除大部分小梯度的样本，缩小分割点的搜索范围，基于直方图算法寻来找最优的分割点，和采用基于叶子节点增长(Leaf-Wise)的策略，在保证模型精度的前提下同时防止过拟合。

#### 2.3.1. 单边梯度采样

单边梯度采样(Gradient-Based One-Side Sampling, GOSS)是 LightGBM 中的一种采样策略，用于在训练过程中提高模型的效率。GOSS 主要通过两种策略来实现训练效率的提高：一种是对样本进行采样，另一种是对特征进行采样。对于样本采样，GOSS 只对梯度值较小的样本进行采样，这些样本的梯度值很小，相当于已经收敛，因此对于模型的训练没有太大的影响，而对于梯度值较大的样本则全部保留，以保证模型的稳定性。对于特征采样，对于某个特征的某个取值，如果其梯度值较小，则对其进行随机采样，而对于梯度值较大的特征取值，则全部保留，以保证模型的稳定性。这样就可以减少计算量，并且能够使得内存占用更小，因为只需要保存符合条件的样本的梯度信息即可。通过单边梯度采样技术，LightGBM 可以在保证模型性能的同时提高训练速度，并且能够应对大规模数据集的训练。

#### 2.3.2. 直方图算法

在传统的梯度提升决策树算法中，每次分裂节点时需要遍历所有的特征进行选择，计算增益后选择最优的特征进行分裂。但是，当数据集非常大时，这个过程会变得非常耗时。而直方图算法通过将数据集按照某个特征进行分块，然后针对每个特征的每个分块计算增益，从而避免了对整个数据集的遍历。这个过程类似于对数据集进行离散化操作，直方图算法通过将原始数据进行离散化，用直方图来近似数据分布的形态，从而加速了决策树的训练过程。直方图算法的增益计算具体如式(6)所示：

$$Gain = \frac{1}{2} \left[ \frac{\left( \sum_{j \in L} g_j \right)^2}{\sum_{j \in L} h_j + \lambda} + \frac{\left( \sum_{j \in R} g_j \right)^2}{\sum_{j \in R} h_j + \lambda} - \frac{\left( \sum_{j \in L \cup R} g_j \right)^2}{\sum_{j \in L \cup R} h_j + \lambda} \right] - \gamma \quad (6)$$

其中， $L$  和  $R$  分别表示两个子区间， $g_j$  和  $h_j$  分别表示第  $j$  个样本的梯度和 Hessian 矩阵， $\lambda$  表示正则化参数， $\gamma$  表示最小分裂增益。

同时 LightGBM 对直方图算法做了一个做差加速的优化，对于每个样本，计算其在当前特征下的梯度和 Hessian 值，然后将其与之前的样本的梯度和 Hessian 值做差，得到当前样本的梯度和 Hessian 差值。这样，我们就可以通过对样本的梯度和 Hessian 差值进行累加，而不是对所有样本进行排序，从而计算出每个分割点的增益。

总的来说，采用了直方图优化的方式，LightGBM 可以快速地计算出每个特征的梯度和样本数，避免了对整个数据集的扫描，从而降低了计算时间和内存占用。

#### 2.3.3. 叶子节点增长策略

在 GBDT 模型中，每次迭代都需要计算出所有样本的负梯度，这对于大规模数据集来说是非常耗时

的。为了解决这个问题，LightGBM 引入了一种新的叶子节点增长策略，采用梯度单边采样和直方图优化的思想，将数据按特征分桶，并按桶计算梯度和样本数，从而提高了计算效率和内存使用效率。它会根据每个叶子的梯度信息和深度限制，选择一个具有最大梯度的样本作为当前叶子的分裂点，并将其分成左右两个子树。这个过程会一直重复，直到达到设定的叶子数目或达到深度限制。其分裂过程如图 5 所示。

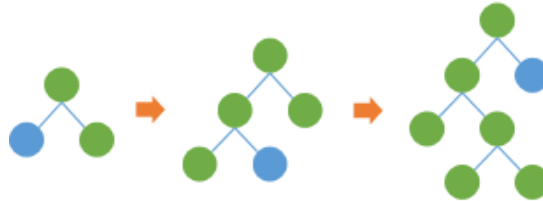


Figure 5. Leaf-wise tree splitting process  
图 5. 叶子节点树分裂过程

这一策略不仅可以高效地处理大规模稀疏数据，而且可以自适应地处理不同特征的重要性差异，从而提高了模型的准确性和可解释性。

此外，LightGBM 的叶子节点增长策略还支持多线程并行计算，可以进一步提高计算效率。同时，还支持特征并行计算和数据并行计算，可以快速处理大规模数据集和高维特征数据。

#### 2.4. 基于 TCN 和 LightGBM 的订单量预测模型

客户订单数据中存在时间相关性特征，而这些特征对于准确快速地预测客户的订单量非常重要。因此，本文提出了一个基于 TCN-LightGBM 的模型。该模型利用 TCN 对客户订单数据进行特征提取，TCN 是一种结合了 CNN 的特征提取能力和 RNN 的时间域建模能力的深度学习网络，它可以有效地捕捉不同时间尺度上的数据依赖关系，并且避免了梯度消失和计算效率低的问题。在提取出时间相关性特征后，本文考虑到实际供应商客户的订单数据规模和复杂性，利用 LightGBM 对提取的特征进行分类或回归。LightGBM 是一种基于决策树的机器学习算法，它可以准确和快速地处理大规模数据，并且具有较强的泛化能力和鲁棒性。通过结合 TCN 和 LightGBM，本文实现了客户订单量的预测，并且在不同场景下进行了验证和评估。基于 TCN-LightGBM 模型的总体框架如图 6 所示。

该框架主要由几个步骤构成，

- 1) 构造特征集：通过整体的竞品和市场消费特征，区域性的竞品和消费特征，以及时序特征等信息，将其构成本文预测模型的输入特征集；
- 2) 特征预处理：对于时序特征，我们使用 one-hot 编码来表示其离散的取值；对于非时序特征，我们使用 Min-Max 归一化来将其数值范围缩放到[0,1]之间。
- 3) 基于 TCN 的特征提取：为了有效地提取输入数据的特征，我们使用基于 TCN 的特征提取网络。TCN 是一种利用一维卷积层和残差连接构建的网络，能够捕捉序列数据中的长期依赖关系。通过使用 TCN，我们不仅能够挖掘特征中的隐藏信息和时间动态，还能够降低特征的维数，减少后续模型的计算量。
- 4) 基于 LightGBM 模型的订单量预测：我们将特征数据划分为训练集、验证集和测试集，分别用于模型的训练、调参和预测。通过多次迭代优化模型参数，我们最终得到了一个较为稳定和准确的 LightGBM 模型，并将其应用于测试集上，对客户的订单量进行了预测。最后，我们使用了一些评价指标，如均方误差、平均绝对误差等，来评估 LightGBM 模型的预测性能，并与其他常用的机器学习模型进行了比较。

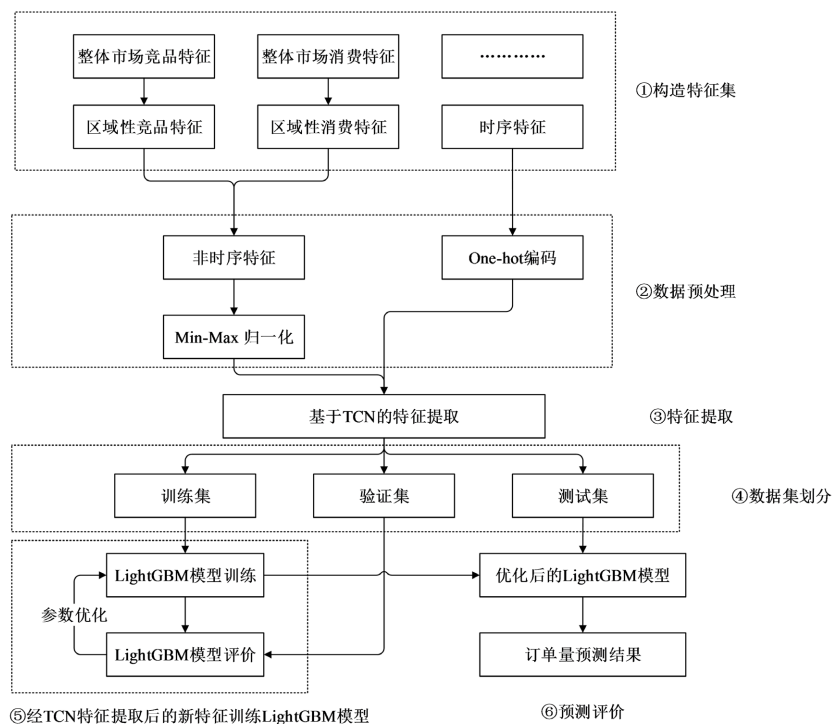


Figure 6. The overall framework of the TCN-LightGBM model  
图 6. TCN-LightGBM 模型的总体框架

### 3. 实验与分析

#### 3.1. 实验数据

为验证本文算法的准确性，通过某快消品生产公司系统中采集 2015 年到 2021 年的所有客户数据、产品基本数据以及订单交易数据，导入大数据平台加工处理后进行分析研究。由于采集数据存在分散、孤立和异构问题，因此需要打通多源、异构数据信息交互壁垒，将大量的、不统一的数据碎片进行全面集成与共享。多源数据差异较大，制造企业一般都是直接将加工后的数据存储到数据库中，其它文件类型数据一般存储在文件服务器上。为直接获取这些数据，本文在导入不同数据源和不同结构数据时采用不同的导入工具，将多源数据集中到数据仓库统一管理，为后续数据处理作准备。

#### 3.2. 实验评价指标

均方根误差(Root Mean Squared Error, RMSE)是一种常用于衡量误差的指标，具体如式(7)所示：

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y^{(i)} - \hat{y}^{(i)})^2} \quad (7)$$

平均绝对值误差(Mean Absolute Error, MAE)是误差绝对值的均值，与 RMSE 相比，MAE 对异常值有更强的鲁棒性，具体如式(8)所示：

$$MAE = \frac{1}{n} \sum_{i=1}^n |y^{(i)} - \hat{y}^{(i)}| \quad (8)$$

平均百分比误差(Mean Absolute Percentage Error, MAPE)是一种无量纲的误差评价指标，可以直观地反映模型的误差水平，具体如式(9)所示：



$$V_{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100\% \quad (9)$$

本文中采用平均平方误差(Mean Square Error, MSE)作为模型训练的损失函数。MSE 函数曲线处处可导, 具有较为稳定的梯度, 具体如式(10)所示:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y^{(i)} - \hat{y}^{(i)})^2 \quad (10)$$

其中,  $y^{(i)}$  与  $\hat{y}^{(i)}$  分别表示实际销售量和预测销售量,  $n$  表示测试数据集大小。MSE、RMSE 或 MAE 值越小, 预测效果越好。

### 3.3. 基准模型

本次实验采用以下基准模型:

**SVR [3]:** SVR 是支持向量回归的缩写, 是一种基于支持向量机(SVM)的回归模型。SVR 的目标是找到一个函数, 使得预测值和真实值之间的误差小于一个给定的阈值, 并且尽可能地平滑。

**LSTM [11]:** LSTM 是长短期记忆网络, 是一种循环神经网络(RNN)的变体。它可以处理长序列数据, 解决梯度消失或爆炸的问题。它由输入门、遗忘门、输出门和记忆单元组成, 可以学习和存储长期依赖关系。

**TCN [13]:** TCN 是时序卷积网络, 是一种卷积神经网络(CNN)的变体。它可以处理不等长的序列数据, 具有因果性和稀疏性。它由多层因果卷积层和残差连接组成, 可以捕获长期依赖关系。

**LightGBM [7]:** LightGBM 是轻量级梯度提升机, 是一种基于树的集成学习算法。它使用基于直方图的算法来加速训练和减少内存消耗。它使用基于叶子的生长策略来增加模型复杂度和减少过拟合。它支持并行学习、类别特征、缺失值处理等特性。

**CNN-LightGBM [15]:** CNN-LightGBM 是一种结合了 CNN 和 LightGBM 的混合模型。它可以利用 CNN 提取特征, 并将其作为 LightGBM 的输入进行分类或回归任务。

### 3.4. 实验结果与分析

为了评估 TCN-LightGBM 模型的性能, 在本文数据集进行实验, 各基准模型参数如表 1 所示。

**Table 1.** Benchmark model parameter table

**表 1.** 各基准模型参数表

模型	参数描述
SVR	核函数为线性核函数, 惩罚因子设为 1
LSTM	隐藏层的数量是 1 和隐藏节点的数量设置为 50
TCN	过滤器的数量是 64, 过滤器的大小是 2 和膨胀系数设置为(1、2、4、8)
LightGBM	树的数量为 790, 最大深度为 3, 叶子节点的数量为 8, 学习率设置为 0.008, 所选择的模型优化器为 Adam, 提升模型为 GBDT
CNN	卷积层的数量是 1, 卷积过滤器的层数是 8, 卷积核的大小是 2, 完全连接层的数量是 2

不同模型的性能对比实验结果如表 2 所示:

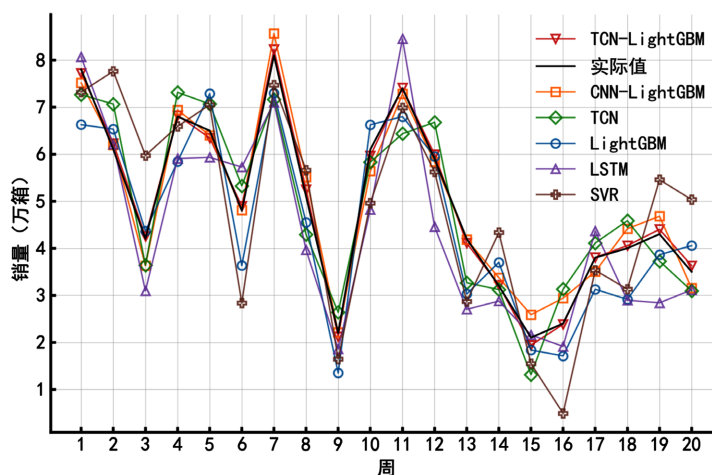
从实验结果可以看出, LightGBM 和 TCN-LightGBM 在三个指标上都取得了最低的数值, 说明它们是较为优秀的模型。特别是 TCN-LightGBM, 在 MAPE 上达到了 2.88%, 远低于其他模型。这说明

**Table 2.** Experimental results of different models  
**表 2.** 不同模型实验结果

评价指标	MAE	MAPE	MSE
SVR	30.56	8.77	42.11
LSTM	26.87	7.46	33.45
TCN	12.51	3.42	16.92
LightGBM	9.81	3.08	14.55
CNN-LightGBM	24.97	7.46	34.08
TCN-LightGBM	9.71	2.88	13.97

TCN-LightGBM 能够更好地捕捉数据的变化趋势，并且对异常值或离群点不敏感。SVR 和 CNN-LightGBM 在三个指标上都取得了较高的数值，说明它们在预测效果上相对较差。可能是因为 SVR 和 CNN-LightGBM 模型在处理非线性、非平稳的时间序列数据时存在一定的局限性。LSTM 则处于中等水平，在三个指标上都没有突出或落后。这可能说明 LSTM 能够一定程度上拟合数据的非线性关系，并且具有一定的鲁棒性。

为验证本文模型的对于预测效果的有效性，本文采用对比实验进行验证，采用 SVR、LSTM、TCN、LightGBM、CNN-LightGBM、TCN-LightGBM 等不同算法预测结果与实际值进行对比如图 7 所示。



**Figure 7.** Comparison of Prediction Results between TCN-LightGBM and Other Algorithms

**图 7.** TCN-LightGBM 与其他算法预测结果对比

从图 7 可以看出，SVR 的预测效果相对较差，主要是因为数据集较大时，随着噪声和异常值的数量也会随着数据集的增大而增加，从而导致 SVR 模型容易出现过拟合，导致模型的性能较差。LSTM 和 TCN 模型预测效果表现较好，主要是他们适用于序列预测问题，处理序列数据表现相对较好。但在处理长序列时，LSTM 的反向传播算法容易出现梯度消失和梯度爆炸问题。梯度消失会导致模型无法学习到长期依赖关系，而梯度爆炸会导致模型不稳定，难以训练。LightGBM 模型的预测效果要优于其他个体模型，因其采用了多种针对大数据集的优化策略，可以显著提高模型训练的效率 and 性能。混合模型中，TCN-LightGBM 的预测效果要优于 CNN-LightGBM，TCN 采用了一些特殊的技巧，如空洞卷积和残差连

接,可以更好地捕获序列数据中的长期依赖关系,从而提高弱学习器的性能,进而提高整个集成模型的性能。

综上所述,本次实验中,基于树的集成学习算法 LightGBM 和时序卷积网络 TCN 以及它们的混合模型 TCN-LightGBM 表现出了较高的预测精度和稳定性。

#### 4. 总结与展望

本文提出了一种基于 TCN 和 LightGBM 的供应商主动维护客户构成模型,旨在解决传统的被动响应式服务无法满足当今竞争激烈市场需求的问题。该模型结合了时间序列分析和机器学习技术,使用 TCN 捕捉客户订单数据的时间依赖性,同时利用 LightGBM 学习非时序的复杂关系特征,对客户需求进行精准预测。在本文数据集上进行实验,通过与 SVR、LightGBM、LSTM、TCN 和 CNN-LightGBM 等模型进行对比,本文模型表现出了较高的预测精度和稳定性。

未来的研究方向包括将该模型应用于更广泛的供应链管理领域,进一步提高模型预测准确性和稳定性,同时探索其他机器学习算法和模型架构对于供应商维护客户构成模型的适用性和效果。此外,结合区块链等新兴技术,可以进一步提高供应商与客户之间的交互效率和数据安全性。这些研究将有助于推动供应链管理的数字化和智能化发展,提高企业的市场竞争力和客户满意度。

#### 基金项目

浙江省重点研发计划(2022C01238)。

#### 参考文献

- [1] 孙延华,张冬杰,曾庆维,等.基于 GBRT 树模型分位数回归预测的 CPF 补货方法[J].软件导刊,2019,18(12):35-39.
- [2] Belgiu, M. and Drăguț, L. (2016) Random Forest in Remote Sensing: A Review of Applications and Future Directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, **114**, 24-31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>
- [3] Smola, A.J. and Schölkopf, B. (2004) A Tutorial on Support Vector Regression. *Statistics and Computing*, **14**, 199-222. <https://doi.org/10.1023/B:STCO.0000035301.49549.88>
- [4] Friedman, J., Hastie, T. and Tibshirani, R. (2000) Additive Logistic Regression: A Statistical View of Boosting (with Discussion and a Rejoinder by the Authors). *The Annals of Statistics*, **28**, 337-407. <https://doi.org/10.1214/aos/1016218223>
- [5] Candelieri, A., Giordani, I., Archetti, F., et al. (2019) Tuning Hyperparameters of a SVM-Based Water Demand Forecasting System through Parallel Global Optimization. *Computers & Operations Research*, **106**, 202-209. <https://doi.org/10.1016/j.cor.2018.01.013>
- [6] Chen, T. and Guestrin, C. (2016) Xgboost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, 13-17 August 2016, 785-794. <https://doi.org/10.1145/2939672.2939785>
- [7] Ke, G., Meng, Q., Finley, T., et al. (2017) Lightgbm: A Highly Efficient Gradient Boosting Decision Tree. *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS' 17)*, Long Beach, 4-9 December 2017, 3149-3157.
- [8] 王华勇,杨超,唐华.基于 LightGBM 改进的 GBDT 短期负荷预测研究[J].自动化仪表,2018,39(9):76-78+82. <https://doi.org/10.16086/j.cnki.issn1000-0380.2018020003>
- [9] 罗蓉.基于机器学习的连锁企业销售预测模型研究[D]:[硕士学位论文].重庆:重庆理工大学,2022. <https://doi.org/10.27753/d.cnki.gcqgx.2022.000227>
- [10] 郭琛.基于深度学习的时间序列组合预测模型研究[D]:[硕士学位论文].南昌:南昌大学,2022. <https://doi.org/10.27232/d.cnki.gnchu.2022.001944>
- [11] Hochreiter, S. and Schmidhuber, J. (1997) Long Short-Term Memory. *Neural Computation*, **9**, 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>

- 
- [12] 杨丽, 吴雨茜, 王俊丽, 等. 循环神经网络研究综述[J]. 计算机应用, 2018, 38(S2): 1-6+26.
- [13] Bai, S., Kolter, J.Z. and Koltun, V. (2018) An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. (Preprint)
- [14] 翟剑锋. 基于 TCN 的时序数据研究[J]. 电子技术与软件工程, 2021(2):196-198.
- [15] 张爱枫, 段新宇, 何泉峰. 基于 CNN 和 LightGBM 的新型风电功率预测模型[J]. 电测与仪表, 2021, 58(11): 121-127. <https://doi.org/10.19753/j.issn1001-1390.2021.11.017>