

关系驱动的基于时间卷积网络的股票走势预测算法

赵 戈, 徐尚志, 李志鹏

同济大学电子与信息工程学院, 上海

收稿日期: 2021年9月22日; 录用日期: 2021年10月21日; 发布日期: 2021年10月28日

摘要

由于股票市场的复杂性和波动性, 股票价格走势预测被认为是一项具有挑战性的任务。基于图卷积网络(Graph Convolution Network, GCN)处理非欧式结构数据的能力, 本文提出了一种股票走势预测算法。该算法通过时间卷积网络(Time Convolution Network, TCN)从股票的历史交易数据中提取特征, 使用图结构对股票之间的关系进行量化, 应用图卷积网络从相关股票中学习局部特征表示来预测股价走势。与以往的研究不同, 本文强调个体特征和多尺度局部特征的融合, 并应用小波包变换对数据进行去噪以提高算法的性能。本文提出的算法在中国大陆沪深300成分股数据集中进行了验证, 在两个指标上取得了优于常见时间序列预测基线模型2%和3%的性能表现。

关键词

图卷积网络, 时间卷积网络, 小波包变换, 股票价格预测

Relation-Driven Stock Trend Prediction Model Based on Time Convolutional Network

Ge Zhao, Shangzhi Xu, Zhipeng Li

School of Electronic and Information Engineering, Tongji University, Shanghai

Received: Sep. 22nd, 2021; accepted: Oct. 21st, 2021; published: Oct. 28th, 2021

Abstract

Stock price movement prediction is considered a challenging task due to the complexity and vola-

文章引用: 赵戈, 徐尚志, 李志鹏. 关系驱动的基于时间卷积网络的股票走势预测算法[J]. 计算机科学与应用, 2021, 11(10): 2555-2567. DOI: 10.12677/csa.2021.1110259

tility of the stock market. Based on the ability of graph convolution network (GCN) to process non-European structured data, this paper proposes a stock trend prediction algorithm. The algorithm extracts features from the historical trading data of stocks through TCN (Time Convolution Network), quantifies the relationship between stocks using graph structure, and uses GCN to learn local feature representation from related stocks to predict the movement of stock price. Different from previous studies, this paper emphasizes the fusion of individual features and multi-scale local features, and applies wavelet packet transform to denoise the data to improve the performance of the algorithm. The algorithm proposed in this paper is verified in the data set of CSI300 constituent stocks in the Chinese mainland. The performance of two benchmark indexes is 2% and 3% better than that of the common time series prediction baseline models.

Keywords

Graph Convolution Network, Time Convolution Network, Wavelet Packet Transform, Stock Price Prediction

Copyright © 2021 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

以股票价格为代表的金融时间序列的预测一直是金融领域关注的焦点。由于股票市场的复杂性，很难凭借主观经验预测股票价格变化情况，相关从业人员一般使用基本面分析法和技术面分析法[1]进行辅助决策。前者通过对宏观经济运行状况、行业结构、上市公司业绩、国家的政策法规等影响证券投资价值以及价格的基本因素进行分析，从而评估证券的实际价值，为投资决策提供理论依据；后者通过统计学方法研究以往价格和交易量的数据，实现对相应证券的分析，进而预测未来证券价格的走向。一种常见的技术面分析法是对股票价格序列进行建模，以预测未来一个交易日的价格变化情况。近些年来随着深度学习技术的不断演进，越来越多的研究者将相关模型和算法应用于技术面分析中，并且取得了相比较统计学模型更好的预测效果[2] [3] [4]。由于历史价格数据波动性较大，为了提高预测准确率，一些研究人员将自然语言处理(Natural Language Processing, NLP)应用到股票预测中，通过分析一些价格数据以外的外部信息，如社交媒体信息[5]和新闻事件[6]，将相关文本信息向量化，结合历史价格数据进行多维度建模分析，对投资者的行为和股价的变化做出更合理的预测。随着图卷积网络[7] (GCN)的提出，深度学习具有了关系推理的能力。传统的深度学习模型在处理时间序列时忽略了实体之间的关系，在金融领域的表现为忽略了股票之间的相关性：例如所在行业、所属概念、公司所在地区以及公司之间的相互持股关系等。由于这些关系的存在，相关股票的价格变动总是有一定的相关性，Chen 等[8]使用图结构整体表示股票市场，用图中的节点标表示公司，用边来量化股票之间的关系，对从历史价格序列中提取的数据特征学习更高维的特征并用于分析预测，取得了优于基线模型的预测精度。

然而目前的工作还存在一些问题：1) GCN 是一种从邻域节点的信息中学习特征表示的算法。以股票价格序列为金融时间序列中包含大量噪声，我们推测来自邻域节点的噪声积累会干扰特征分析。2) 单一尺度的分析无法准确捕捉节点的实际特征。K 层的 GCN 网络的节点表示强调 K 阶邻域的局部信息，因此同一连通分量内的节点表示趋于平滑化，不利于每个节点的具体分析。

针对上述问题，本文提出了一种基于 TCN 和 GCN 的股票价格预测模型，该模型从股票过去 D 天的

历史价格和股票之间的关系中学习高阶特征表示，用于第 $D + 1$ 天的股价走势分类预测。具体来说，我们选取沪深 300 成分股作为分析对象，首先对所有成分股的历史价格序列进行小波包变换以消除噪声。第二步，我们使用在时间序列分析上比 RNN (Recurrent Neural Network)、LSTM (Long Short-Term Memory) 系列深度学习模型表现更好的 TCN 模型来完成股票价格序列的特征向量的提取。第三步，我们使用本文设计的多尺度 GCN 模型从获得的特征中学习股票图的局部特征以进行走势预测。这项工作侧重于特征的多尺度分析，所提出的模型聚合了从历史数据中获得的原始特征和不同 GCN 层的输出，以保证在综合分析不同尺度的局部特征的同时不会丢失节点的个体特征。我们在沪深 300 成分股中进行了相关实验，并采用滑动时间窗法验证了结果。在大约 1100 天的数据中，与常用的时间序列基线模型相比，我们模型的分类预测精度和 F1 指数平均提高了 2%~3%。

本文的其余部分安排如下。第 2 节介绍了本文提出的算法的结构和技术细节。实验设置和结果分析在第 3 节中显示。在第 4 节进行了实验结果的展示和分析。在最后一节中进行了工作总结和对未来工作的展望。

2. 关系驱动的基于时间卷积网络的股票走势预测算法设计

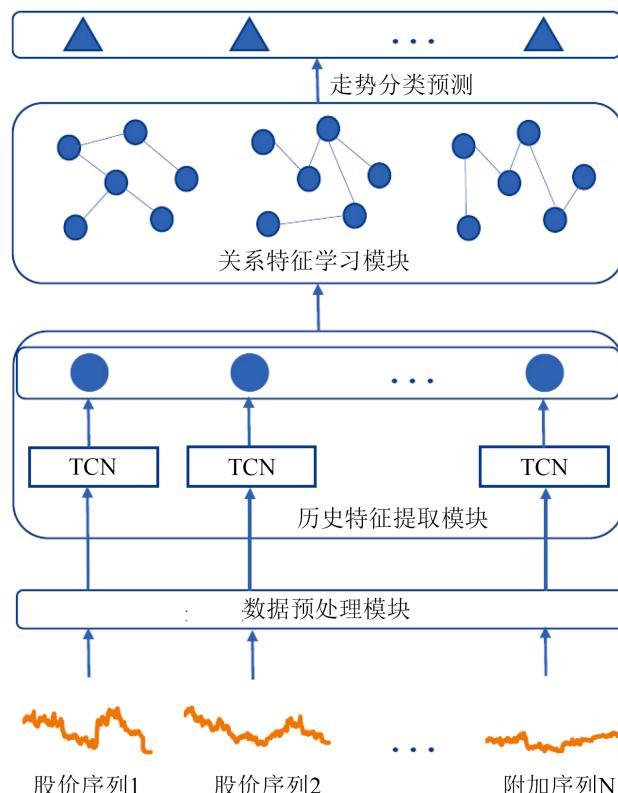


Figure 1. Algorithm framework
图 1. 算法框架

2.1. 预测算法框架

使用数学模型进行股价预测一般有两种方式：将预测转化为回归任务或者分类任务。前者预测具体的股票价格，而后者预测股票的价格走势。本文中采用后者，并将未来一天的股价走势 Movement 分为三个场景：[涨(UP)，小幅波动(STAY)，跌(DOWN)]如式(1)中所定义， P_t 代表第 t 天收盘价相较于前一天

6 的涨幅:

$$\text{Movement} = \begin{cases} \text{UP} & \text{if } P_t \geq 0.5, \\ \text{DOWN} & \text{if } P_t \leq -0.5, \\ \text{STAY} & \text{else.} \end{cases} \quad (1)$$

本文提出了一个集成 TCN 和 GCN 架构的用于股票价格走势预测的深度学习算法模型, 模型的结构如图 1 所示, 该模型包括三个模块: 1) 数据预处理模块。通过小波包变换对股价时间序列进行分解、去噪和重建; 2) 历史特征提取模块。通过 TCN 网络从去噪序列数据中获取特征向量; 3) 关系特征学习模块。使用本文设计的 Multi-GCN 网络学习融合了个体特征和局部结构特征的特征向量, 用于最终的分类预测。各模块中的技术细节如下。

2.2. 数据预处理模块

在本模块中, 我们使用小波包变换对原始股价时间序列进行去噪。小波包变换是小波变换的一种改进形式, 广泛应用于信号处理领域[9]。由于小波变换对时间序列的多尺度分解能力, 它也常被应用于金融时间序列的去噪[10]。小波包变换从多个时域和频域尺度分解金融时间序列, 其中, 噪声分量被认为具有比信号分量更低的小波包系数, 因此我们可以设置一个阈值来过滤噪声分量。小波包变换的基本步骤是 1) 选择小波基; 2) 确定分解层数; 3) 计算阈值; 4) 选择阈值函数对序列进行过滤。本文选用具有良好的正交性和紧凑性 Daubechies 小波系中的 DB4 作为小波基来完成小波包分解。小波包去噪的关键在于计算阈值的方法和阈值函数的选择。常见的计算方法包括固定阈值估计、极值阈值估计、无偏似然估计和启发式估计等。对于金融时间序列, 信号分量分布在每个频带中, 为了保护信号在去噪过程中不被破坏, 我们选择较为保守的无偏似然估计方法作为阈值计算方法。无偏似然估计方法首先取信号中每个元素的绝对值, 并将它们从小到大排列并取每个元素的平方(公式(2))。然后按照公式(3)设置每个元素的风险值。最后将风险值最低的元素设置为阈值(公式(4))。

$$f(k) = \left(\text{sort}(s)^2 \right), (k = 0, 1, 2, \dots, N-1) \quad (2)$$

$$Risk(k) = \left[N - 2k + \sum_{i=1}^k f(i) + (N-k)f(N-k) \right] / N \quad (3)$$

$$Th = \sqrt{\text{argmin}Risk(k)} \quad (4)$$

本文采用了郭晓霞等[11]提出的改进阈值函数(公式(5)), 克服了硬阈值函数的不连续性, 同时减少了使用软阈值函数造成的重构信号与原始信号的偏差。

$$\text{Threshold}(x, Th) = \begin{cases} \text{sgn}(x)(|x| - 0.5Th) & x \geq Th, \\ 0 & x < Th. \end{cases} \quad (5)$$

在该模块中首先对数据集中每只股票过去 D 天的价格数据进行小波包变换去噪。股票价格时间序列被分解为不同频段的分量。 s_0 代表最低频率分量, 它被认为是与原始信号具有相似趋势的近似分量。其他分量被认为是细节分量, n 表示分解的层数, 在本文中根据序列的长度设置为 3。分解完成后对细节分量进行去噪, $T(\cdot)$ 指的是阈值过滤的过程。对已经去噪后的分量进行重构, 得到去噪信号 S_{out} , 如公式(6)所示。

$$S_{out} = s_0 + T(s_1) + T(s_2) + \dots + T(s_{2^n-1}) \quad (6)$$

最后对去噪后的股票价格序列进行归一化(公式(7))完成数据预处理工作, 其中 \tilde{x} 是序列的均值, σ 是

序列的方差。

$$x^* = \frac{x - \tilde{x}}{\sigma} \quad (7)$$

2.3. 历史特征提取模块

在本模块中，我们采用 TCN 网络[12]从过去 D 天的股票价格时间序列中提取特征向量，用于后续进一步分析。TCN 是一种一维全卷积网络[13]，以填充 0 的形式保证每个隐藏层的长度与输入层的长度相同。如图 2 左所示，TCN 网络结构中应用了三个关键设计：因果卷积、扩张卷积和残差连接，这赋予了 TCN 强大的时间序列分析能力。下面，我们将介绍这些设计及其在这项工作中的应用。

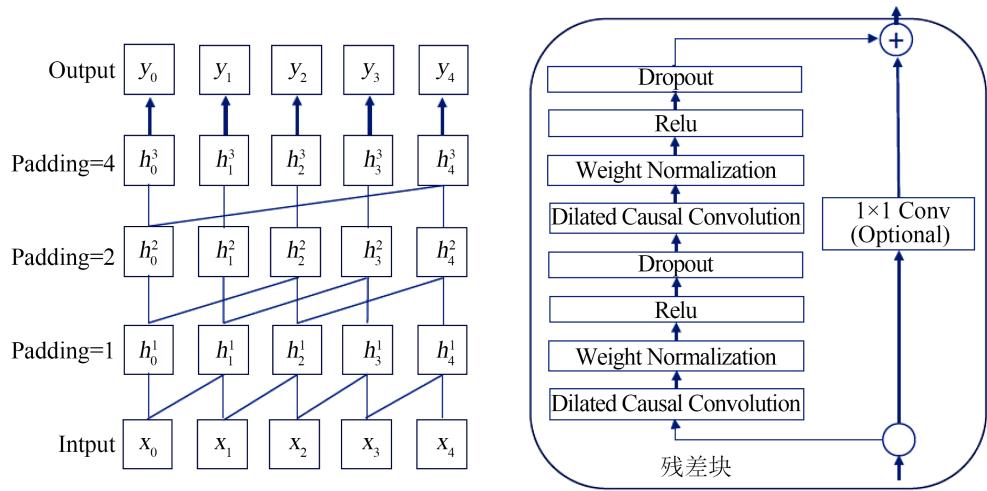


Figure 2. Structure of TCN (left) and Residual Block (right)
图 2. TCN 结构(左)和残差块(右)

2.3.1. 因果卷积

因果卷积最早是在 WaveNet [14]中提出的，它的主要设计思想是确保卷积网络在分析时间序列时不会从序列中的未来数据中获取信息，如图 3 左所示，网络在时刻 t 的输出仅取决于时刻 t 和时刻 t 之前网络的输入。由于因果卷积没有经常被用于处理时间序列的 LSTM 网络的循环连接结构，因此因果卷积的处理速度通常更快。

2.3.2. 扩张卷积

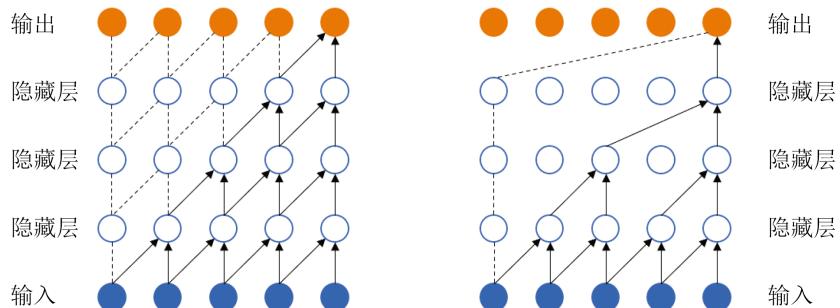


Figure 3. Causal convolution (left) and extended convolution (right)
图 3. 因果卷积(左)和扩张卷积(右)

由于因果卷积的感受野与网络层数线性相关,为了降低计算成本,TCN采用了扩张卷积的形式(图3右),跳过部分输入单元进行间隔采样,使感受野随着层数指数级增大[15]。对于一维序列输入 $X \in \mathbb{R}^n$ 和卷积核 f ,对序列元素 x_i 的扩张卷积运算 F 定义为公式(8),其中 d 是扩张因子(采样率),数量上为 2 的阶乘, k 是卷积核尺寸。

$$F(x_i) = (X *_d f)(x_i) = \sum_{i=0}^{k-1} f(i) \cdot X_{x_i - d \cdot i} \quad (8)$$

2.3.3. 残差连接

残差连接结构最早在 ResNet 网络被提出[16],一般用于在网络中跨层传递信息,解决网络加深带来的梯度消失问题。常见的方法是将输入张量处理后加入到残差块的输出张量中,由于输入和输出张量可能具有不同的尺寸,因此 TCN 应用了额外的 1×1 卷积层以保持输入和输出张量的形状相同。TCN 使用残差块代替卷积层,如图 2 右所示,每个残差块包含 2 组扩张卷积层和非线性连接层 ReLU [17],此外每组内还添加了权重归一化层[18]和 dropout [19] 层以进行正则化。残差块中的计算过程可以用公式(9)表示。

$$Z_t^l = R(Z_t^{l-1}) + W_{1 \times 1} Z_t^{l-1} \quad (9)$$

其中 $R(\cdot)$ 表示残差块中的运算处理过程, $W_{1 \times 1}$ 表示 1×1 卷积。 Z_t^l 是 TCN 网络第 l 层残差块在时刻 t 的输出。

2.4. 关系特征学习模块

在本模块中我们主要完成了两项工作:1)根据股票的不同关系建立三种股票的关系图;2)使用多尺度 GCN 网络获取结构特征,即捕捉同行业股票、同概念股票和相互持股的公司的整体价格走势特征,并进行最终走势的预测。下面,我们将介绍该模块的细节。

2.4.1. 实体关系图构建

通常的,对于一张图 G ,图包含一组节点以及它们的边。在这里我们用图中的点表示一支股票,用边表示股票直接的关系。我们尝试从行业、概念和相互持股等常见股票分类中对股票之间的关系进行量化,并完成三种关系图的构建。对于同行业的股票,我们将这两只股票所代表的节点设置为连通,权重为 1。概念在中国股市中经常被用来描述股票涉及的业务范围或具有的一种特点。例如,5G 概念股包括中国联通,一般认为 5G 的相关政策和信息会对其股价的变动产生一定的影响。对于两支具有相同概念的股票,我们设置关系权重为共同概念的数量除以每只股票具有的概念数量。对于有持股关系的公司,我们简单地以持股比例作为权重值。

2.4.2. 多尺度 GCN 网络

对于历史特征提取模块获得的特征张量,我们将其置于关系图的节点中作为每个节点的表示。GCN 网络定义了一个具有分层传播规则的图卷积网络,对于第 l 层网络的节点特征 $H^{(l)}$,该规则如公式(10)所示。

$$H^{(l+1)} A = \sigma \left(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{\frac{1}{2}} H^{(l)} W^{(l)} \right) \quad (10)$$

其中 $\tilde{A} = A + I$, A 是图的邻接矩阵, I 是单位矩阵, \tilde{D} 是 \tilde{A} 的度矩阵, σ 为非线性激活函数 ReLU, W 为特征变换矩阵。由于 GCN 本身的结构特性,随着网络的加深,邻域节点的特征趋于收敛,节点原始的个体特征会发生丢失。为了避免这种现象,在原始的 GCN 的基础上,如图 4 所示,我们定义了一个用于多种关系图的多尺度 GCN 网络模型,可以在不失去原有个体特征的同时有效地传播结构图上的信息。我

们将 GCN 层的输出与前层的输出拼接起来，并输入到下一层的输入中，这样可以对来自不同区域范围的局部结构特征进行多尺度分析。多尺度 GCN 网络传播过程如公式(11)所示。其中 $i=1,2,3$ 指由三种不同关系构建的邻接矩阵， t 为当前时刻。

$$H_{i,t}^{(l+1)} = concat\left(\sigma\left(\widetilde{D}_i^{-\frac{1}{2}}\widetilde{A}_i\widetilde{D}_i^{\frac{1}{2}}H_{i,t}^{(l)}W_i^{(l)}\right), H_{i,t}^{(l)}\right) \quad (11)$$

对于从三个图中获得的复合特征 $H_{i,t}^{(l)}$ ，我们使用注意力机制将其融合。如公式(12)和公式(13)所示，将三个图的输出特征转化为向量，然后计算每个图的权重。

$$e_i = W_i H_{i,t}^{(l)} + b_i \quad (12)$$

$$a_i = \frac{\exp(e_i)}{\sum_{k=1}^3 \exp(e_k)} \quad (13)$$

$$F_{at,t} = \sum_{k=1}^3 a_i H_{i,t}^{(l)} \quad (14)$$

其中 e_i 表示用于计算权重的每个图的向量， a_i 表示每个图的权重。最后，如公式(14)所示，对三个图的特征进行加权平均，得到算法的输出向量 $F_{at,t} \in \mathbb{R}^{N \times F}$ ，用于股票价格走势的分类预测，其中 N 为股票的数目， F 为特征的维度。

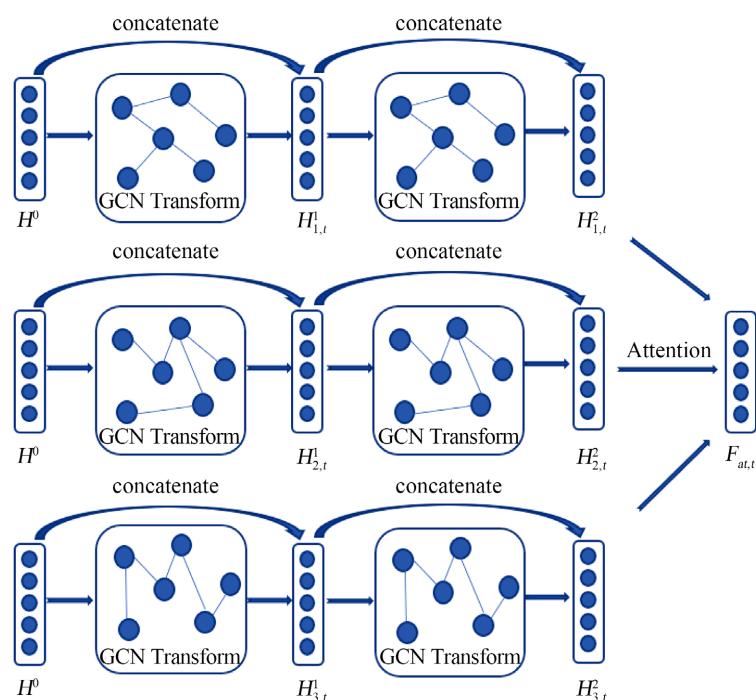


Figure 4. Multi-GCN Structure
图4. 多尺度 GCN 网络结构

3. 数据集和实验设置

3.1. 实验数据集

我们选取沪深 300 成分股 2016 年 6 月 1 日至 2019 年 7 月 5 日的交易数据制作数据集，该数据来自

股票数据接口 Tushare (<https://tushare.pro/>)。如表 1 所示，每行数据包含 6 种交易数据，分别是该股票在当天的开盘价，收盘价，最高价，最低价，交易量，交易额。

Table 1. Data sample of CSI 300 constituent stocks
表 1. 沪深 300 成分股数据样本

股票代码	日期	开盘价	最高价	最低价	收盘价	交易量	交易额
000001.SZ	2015/06/01	15.33	15.98	15.19	15.90	2,158,362	3,365,459
000001.SZ	2015/06/02	15.89	15.90	15.52	15.77	1,904,822	2,991,564
.....
603993.SH	2019/07/05	4.02	4.03	4.03	4.01	315,105	126,263

此外，如表 2 所示，我们还从该 API 中获取了所选股票的一些基本信息用于构建实体的关系图，这些基本信息包括该股票所属行业、概念、公司所在地以及公司间持股情况。

Table 2. Basic information of CSI 300 constituent stocks
表 2. 沪深 300 成分股基本信息

股票代码	公司名称	行业	城市	概念
000001.SZ	平安银行	金融	深圳	深圳本地股、区块链.....
.....
603993.SH	洛阳钼业	稀有金属	洛阳	稀有资源、小金属.....

3.2. 滑动窗口法

为了保证验证实时性，我们选择滑动窗口法对算法进行验证。我们将数据集中的数据划分为八个区间(区间的开始和结束时间如表 3 所示)。每个窗口区间包含 200 天的训练数据、50 天的验证数据和 50 天的测试数据。滑动窗口的步长为 100 天。与直接划分验证集和训练集相比，滚动窗口法时间跨度较小，可以保证训练出来的模型是最。同时，不同时间段的验证也可以对算法性能做出更可靠的评价。

Table 3. Start and end time of sliding windows
表 3. 滑动窗口的起止时间

	Window0	Window1	Window2	Window3	Window4	Window5	Window6	Window7
开始	2015-06-01	2015-10-29	2016-03-25	2016-08-18	2017-01-17	2017-06-20	2017-11-14	2018-04-13
结束	2016-08-17	2017-01-16	2017-06-19	2017-11-13	2018-04-12	2018-09-04	2019-02-01	2019-07-16

3.3. 实验设置

本文使用 Pytorch 框架用于算法模型的训练，选择 Adam 优化器，设置学习率为 0.005，epoch 为 1000。根据原始输入数据的维度和训练成本，我们设置 GCN 的隐藏层单元数为 32。对于模型的一些超参设置，我们进行了一些对比实验来比较：选择滑动窗口的训练数据跨度为 200 天，并设置了 100、150、250 作为对照组；选择被预测目的前 7 天的交易数据作为输入样本，并设置 5、10、15 天作为对照组。本文选择交叉熵作为训练的损失函数。此外，我们还设置了一些常用的基线算法与本文提出的算法进行比较，包括 1) MLP; 2) CNN; 3) RNN; 4) LSTM; 5) GRU; 6) TCN; 7) GCN; 8) L-GCN; (LSTM + GCN); 9) G-GCN (GRU + GCN); 10) T-GCN (TCN + Multi-GCN)。

本文中使用准确率(Accuracy)和 F1 指标来衡量模型的性能。前者衡量模型的预测能力, F1 衡量模型在不同类别中的综合。TP、TN、FP 和 FN 分别表示:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (15)$$

$$\text{F1} = \frac{2 \times \text{TP}}{2 \times \text{TP} + \text{FP} + \text{FN}} \quad (16)$$

其中 TP 指正确分类的正样本数量, TN 指正确分类的负样本数量, FP 指未正确分类的正样本数量, FN 指未正确分类的负样本数量。

4. 实验结果与分析

4.1. 去噪效果分析

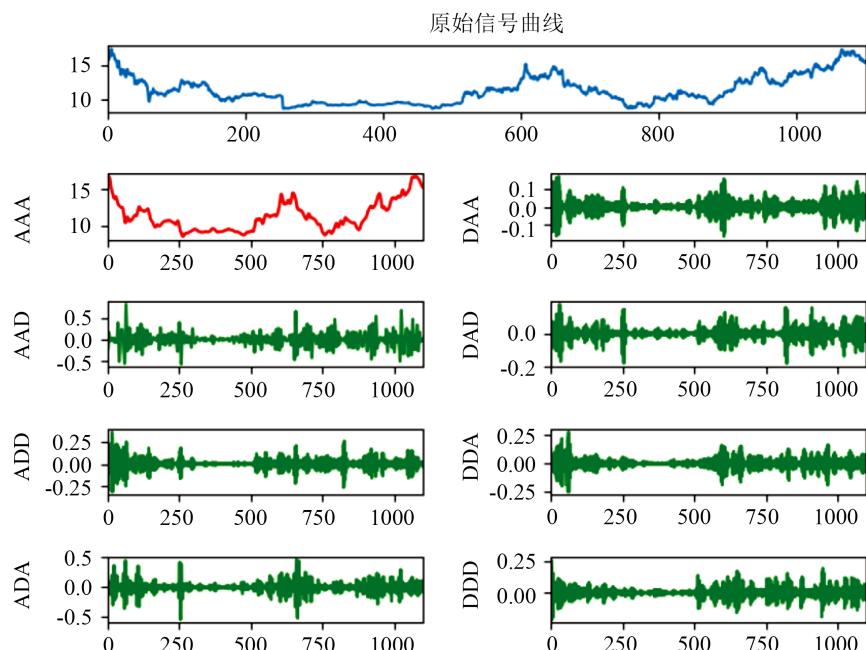


Figure 5. Component curves of original signal curve after wavelet packet decomposition
图 5. 原始信号曲线小波包分解后的分量曲线

本文中使用三层小波包变换, 如图 5 所示, 将原始信号曲线分解为八个不同频带的分量。其中 A (Approximate) 代表每层的近似分量, D (Detail) 代表每层的细节分量。将分量曲线与原始信号曲线进行比较, 可以观察到其中 AAA 分量包含与原始信号相同的走势, 即该分量基本不包含噪声, 应予以保留, 而其它分量包含着一定的噪声, 应予以过滤。在对分量进行去噪和重构后, 如图 6 所示, 我们得到去噪重组后的信号和偏差曲线(表示原始信号曲线和去噪信号曲线在每个时刻的偏差值)。过滤后的曲线在保留曲线整体走势的同时, 明显消除了小的噪声波动, 这更有利于我们进一步的特征提取和分析。由图 6 可以观察到, 当原始曲线位于上升或下降段(0~200 和 500~750 等)时, 噪声量(偏差值)显著增加, 而在稳定段(250~500 等), 噪声可以忽略不计, 这种现象实际上是由投资者情绪变化造成的[20], 当股票价格经历明显变化时, 投资者会更频繁地进行交易以获得利润。然而由于投资者的行为的复杂性, 反映在价格曲线上为不规则的抖动, 同时也会对别的投资者的行为造成影响。此外, 不同频带中的抖动和持续时间是不

同的,这意味着不同频带中分量的噪声幅度和时间分布是不同的。因此,有必要对噪声进行多尺度去噪。去噪是否可以改善算法的性能,我们在下节的实验中进行了验证。

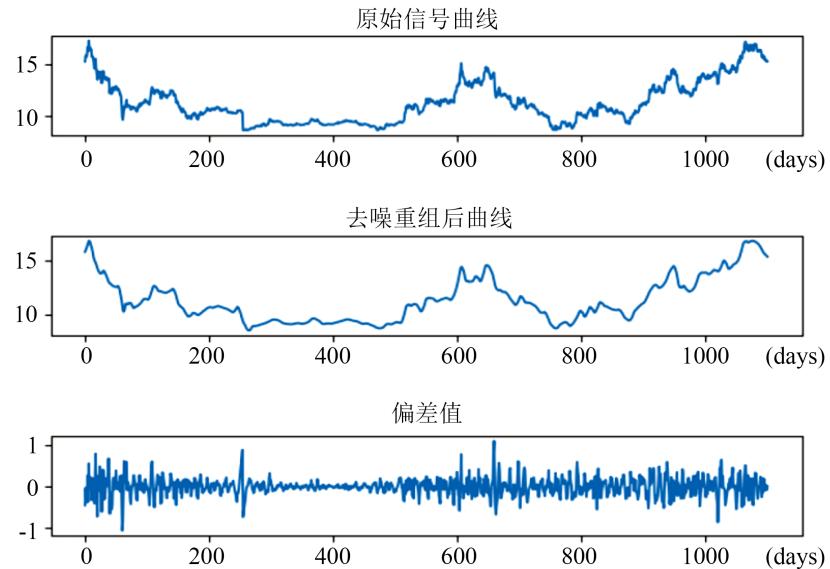


Figure 6. Comparison between original signal curve and denoised reconstructed curve
图 6. 原始信号曲线和去噪重构后曲线对比

4.2. 预测性能分析

Table 4. Comparison of prediction accuracy of algorithm

表 4. 算法预测准确率对比

模型	Window0	Window1	Window2	Window3	Window4	Window5	Window6	Window7	平均
MLP	0.3443	0.3558	0.3517	0.3818	0.3872	0.3638	0.3513	0.3411	0.3596
CNN	0.3622	0.3502	0.3396	0.3917	0.3839	0.3547	0.3564	0.3672	0.3632
RNN	0.3669	0.3813	0.3618	0.3965	0.3813	0.3662	0.3672	0.3527	0.3717
LSTM	0.3906	0.3638	0.3812	0.4015	0.4012	0.3923	0.3987	0.3664	0.3869
GRU	0.3959	0.3614	0.3701	0.4166	0.4052	0.3943	0.3769	0.3734	0.3867
TCN	0.4011	0.3712	0.3801	0.4212	0.4114	0.4249	0.3895	0.3667	0.3957
GCN	0.3712	0.3688	0.3742	0.3929	0.3984	0.4023	0.3752	0.3891	0.3840
L-GCN	0.3863	0.4073	0.4154	0.4163	0.4287	0.4064	0.3989	0.3861	0.4056
G-GCN	0.3817	0.3953	0.4021	0.4136	0.4238	0.4126	0.3919	0.4186	0.4049
T-GCN	0.4181	0.4267	0.4126	0.4285	0.4397	0.4274	0.4125	0.4168	0.4239

根据表 4 和表 5 呈现的实验结果和引言中提出的问题,我们作如下分析:

1) 在表 4 和表 5 中,我们比较了在股票走势分类预测任务中使用和不使用关系特征(即是否使用到图网络)的算法的性能。除原始 GCN 外,GCN 结合常见的时间序列特征提取网络(LSTM、TCN 等)使用,对预测精度都有一定的提升。与基线模型相比,本文提出的组合模型 T-GCN 表现出最好的平均性能,两个指标的平均结果提高了约 2% 和 3%。对于表 4 所示的准确率的指标,本文设计的组合模型在 8 个滑动窗口中的 6 个表现最好,对于表 5 所示的 F1 指标,它在 8 个滑动窗口中的 4 个中表现最好。GCN 算法

完成局部特征的平均并纠正一些错误的走势预测。因此，整体平均精度得到了提高。与之同时，一些股票的走势由于个体原因与局部走势不一致，GCN 算法在计算的过程中模糊了这部分股票的个体走势，这就解释了为什么一些 GCN 组合算法的 F1 值与 TCN 等基础模型相比有所下降。然而，本文提出的 T-GCN 通过特征融合兼顾了个体特征和多尺度局部特征，从实验结果上验证了多尺度结构可以提高算法的性能，有效防止了个体特征的丢失，增强了算法的泛化能力。

Table 5. Comparison of F1 index of algorithm prediction results**表 5.** 算法预测结果 F1 指标对比

模型	Window0	Window1	Window2	Window3	Window4	Window5	Window6	Window7	平均
MLP	0.2871	0.3668	0.2804	0.3979	0.4098	0.2247	0.2501	0.2949	0.3139
CNN	0.2598	0.3409	0.2622	0.3749	0.3819	0.2047	0.2121	0.2929	0.2911
RNN	0.2891	0.3763	0.2817	0.3654	0.3849	0.3642	0.3037	0.3076	0.3341
LSTM	0.2993	0.3362	0.3784	0.3751	0.3672	0.3427	0.3426	0.3089	0.3438
GRU	0.2806	0.3521	0.3676	0.3495	0.3660	0.3341	0.3972	0.3062	0.3441
TCN	0.3357	0.3671	0.3696	0.3878	0.3543	0.3801	0.3566	0.3244	0.3594
GCN	0.2378	0.2924	0.3027	0.3133	0.3344	0.2719	0.2402	0.2672	0.2824
L-GCN	0.2554	0.3438	0.3398	0.3629	0.3762	0.3279	0.3221	0.3237	0.3314
G-GCN	0.2913	0.3575	0.3254	0.3733	0.3866	0.3329	0.3352	0.3472	0.3436
T-GCN	0.3013	0.4094	0.4050	0.4110	0.4361	0.3775	0.3945	0.2741	0.3761

Table 6. Performance comparison of T-GCN using different denoising threshold functions**表 6.** T-GCN 使用不同的去噪阈值函数的性能对比

	不去噪	软阈值	硬阈值	改进阈值
Accuracy	0.3945	0.4132	0.4091	0.4239
F1	0.3438	0.3715	0.3639	0.3761

2) 从表 6 中可以观察到，对比去噪和不去噪的数据，经过小波包变换去噪后的数据使算法性能得到了明显的提升，同时也验证了去噪可以有效提升图卷积模型在股价走势预测领域的性能。与其他阈值函数相比，本文使用的改进的阈值函数展示了最佳的性能。它保留了更多的走势信息，处理后的数据更有利于算法的分析预测。

Table 7. Performance comparison of sliding window under different training data time span**表 7.** 滑动窗口不同训练数据时长跨度下的性能对比

	100	150	200	250
Accuracy	0.3762	0.3904	0.4239	0.4158
F1	0.3461	0.3565	0.3761	0.3745

Table 8. Performance comparison of input data under different time spans**表 8.** 输入数据不同时间跨度下的性能对比

	5	7	10	15
Accuracy	0.3935	0.4239	0.4132	0.3944
F1	0.3541	0.3761	0.3680	0.3656

3) 在表7和表8中, 我们对一些超参进行了比较。从表7可以看出, 对于滑动窗口的训练数据时长跨度, 在200天时算法的性能达到最佳; 输入数据的时间跨度对模型的性能也有影响, 与5、10、15的时间跨度相比, 7天的输入数据展示了最佳的性能。我们可以分析得到结论: 在瞬息万变的股票市场中, 模型学到的规律是有时效性的, 在实际预测中需要针对具体情境进行参数的调整。

5. 总结与展望

本文提出了一种新的关系驱动的股票走势预测算法, 该算法用图结构表示股票实体之间的关系, 通过小波包变换去噪后使用TCN和改进的GCN网络提取不同尺度的特征。通过实验对比可以得出结论, 用于多尺度分析的特征拼接组合结构可以有效减少GCN网络传播造成的单个特征的损失, 提高算法的性能。此外, 在使用小波包变换进行数据去噪后, 图网络算法的特征提取性能可以得到显著提高, 这也可以应用于类似的关系型时间序列分析预测任务。在未来的工作中, 我们将探索如何更好地完成股票之间关系的表示, 以及进一步优化个体特征和局部特征的组合, 提高算法在实际场景中的预测精度。

基金项目

这项工作得到了国家自然科学基金的资助, 资助号为61773290和71571107。

参考文献

- [1] 张玉婧. 股市基本面分析与技术面分析结合方法探讨[J]. 商业文化(上半月), 2011(10): 201.
- [2] Fischer, T. and Krauss, C. (2017) Deep Learning with Long Short-Term Memory Networks for Financial Market Predictions. *European Journal of Operational Research*, **270**, 654-669. <https://doi.org/10.1016/j.ejor.2017.11.054>
- [3] Zhang, R., Yuan, Z. and Shao, X. (2018) A New Combined CNN-RNN Model for Sector Stock Price Analysis. 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC), Tokyo, 23-27 July 2018, 546-551. <https://doi.org/10.1109/COMPSAC.2018.10292>
- [4] 赵红蕊, 薛雷. 基于LSTM-CNN-CBAM模型的股票预测研究[J]. 计算机工程与应用, 2021, 57(3): 203-207.
- [5] Xu, Y. and Cohen, S.B. (2018) Stock Movement Prediction from Tweets and Historical Prices. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, Vol. 1: Long Papers, Melbourne, 15-20 July 2018, 1970-1979. <https://doi.org/10.18653/v1/P18-1183>
- [6] 张梦吉, 杜婉钰, 郑楠. 引入新闻短文本的个股走势预测模型[J]. 数据分析与知识发现, 2019, 3(5): 11-18.
- [7] Kipf, T.N. and Welling, M. (2016) Semi-Supervised Classification with Graph Convolutional Networks. arXiv:1609.02907.
- [8] Chen, Y. and Wei, Z. and Huang, X. (2018) Incorporating Corporation Relationship via Graph Convolutional Neural Networks for Stock Price Prediction. *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, Torino, 22-26 October 2018, 1655-1658. <https://doi.org/10.1145/3269206.3269269>
- [9] Walczak, B. and Massart, D.L. (1997) Noise Suppression and Signal Compression Using the Wavelet Packet Transform. *Chemometrics and Intelligent Laboratory Systems*, **36**, 81-94. [https://doi.org/10.1016/S0169-7439\(96\)00077-9](https://doi.org/10.1016/S0169-7439(96)00077-9)
- [10] Ramsey, J.B. (1999) The Contribution of Wavelets to the Analysis of Economic and Financial Data. *Philosophical Transactions of the Royal Society A: Mathematical Physical & Engineering Sciences*, **357**, 1593-2606. <https://doi.org/10.1098/rsta.1999.0450>
- [11] 郭晓霞, 杨慧中. 小波去噪中软硬阈值的一种改良折衷法[J]. 智能系统学报, 2008, 3(3): 222-225.
- [12] Bai, S., Kolter, J.Z. and Koltun, V. (2018) An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling. arXiv preprint arXiv:1803.01271..
- [13] Long, J., Shelhamer, E. and Darrell, T. (2015) Fully Convolutional Networks for Semantic Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**, 640-651. <https://doi.org/10.1109/TPAMI.2016.2572683>
- [14] Van Den Oord, A., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., Kalchbrenner, N., Senior, A. and Kavukcuoglu, K. (2016) Wavenet: A Generative Model for Raw Audio. arXiv preprint arXiv:1609.03499.
- [15] Yu, F. and Koltun, V. (2016) Multi-Scale Context Aggregation by Dilated Convolutions. *International Conference on*

- Learning Representations 2016*, San Juan, 2-4 May 2016, 1-13.
- [16] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 27-30 June 2016, 770-778.
<https://doi.org/10.1109/CVPR.2016.90>
 - [17] Glorot, X., Bordes, A. and Bengio, Y. (2011) Deep Sparse Rectifier Neural Networks. *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics (AISTATS)*, Fort Lauderdale, 11-13 April 2011, 315-323.
 - [18] Salimans, T. and Kingma, D.P. (2016) Weight Normalization: A Simple Reparameterization to Accelerate Training of Deep Neural Networks. *Advances in Neural Information Processing Systems*, **29**, 901-909.
 - [19] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R. (2014) Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, **15**, 1929-1958.
 - [20] 何柏彬. 投资者情绪对我国股票价格波动的影响研究[D]: [硕士学位论文]. 杭州: 浙江大学, 2021.