

# 基于改进 Transformer 时序算法的 区域经济预测

刘海宏<sup>1</sup>, 刘 敏<sup>2</sup>, 朱岸青<sup>3</sup>

(1.广州南洋理工职业学院经济管理学院, 广东 广州 510900)

(2.西南石油大学计算机与软件学院, 四川 南充 637001)

(3.暨南大学管理学院, 广东 广州 510000)

[摘要] 现有的区域经济预测模型指标变量存在冗余,且忽略了行业、区域等静态变量对预测结果的影响,导致预测效率不高.针对上述问题,提出了基于改进 Transformer 时序算法的区域经济预测模型.首先利用 Copula 函数对传统 Transformer 模型进行优化(Coformer);其次选取区域经济的影响指标,对其进行主成分分析,去除冗余信息;然后将降维后的指标变量和静态变量作为 Coformer 的输入,对变量进行编码,并通过多头注意力机制增强重要信息,最后用解码器对编码的变量解码,利用 Softmax 输出区域历年生产总值序列的预测结果.实验结果表明,所提模型的预测准确率为 0.908,比另外三种模型分别提高了 15.9%、12.3% 和 6.7%,表现出了优异的预测性能.

[关键词] 区域经济预测, Transformer, Copula 函数, 主成分分析, 多头注意力

[中图分类号] TP391 [文献标志码] A [文章编号] 1001-4616(2024)04-0118-08

## Regional Economic Forecasting Based on Improved Transformer Sequence Algorithm

Liu Haihong<sup>1</sup>, Liu Min<sup>2</sup>, Zhu Anqing<sup>3</sup>

(1.School of Economics and Management, Guangzhou Nanyang Polytechnic College, Guangzhou 510900, China)

(2.School of Computer and Software, Southwest Petroleum University, Nanchong 637001, China)

(3.School of Management, Jinan University, Guangzhou 510000, China)

**Abstract:** The existing regional economic forecasting model indicators are redundant and ignore the impact of static variables such as industry and region on the forecast results, leading to low forecasting efficiency. In response to the above issues, a regional economic forecasting model based on an improved Transformer time series algorithm is suggested. Firstly, the traditional Transformer model is optimized using a Copula function(Coformer); secondly, the impact indicators of the regional economy are selected, and principal component analysis is performed on them to remove redundant information; then, the reduced-dimensional indicator variables and static variables are used as the input of the Coformer, and the variables are encoded. Finally, the decoder decodes the encoded variables and uses Softmax to output the prediction results of the regional gross domestic product series over the years. The experimental outcome indicates that the designed model has an accuracy of 0.908, which is 15.9%, 12.3% and 6.7% higher than the other three models, respectively, demonstrating excellent predictive performance.

**Key words:** regional economic forecasting, Transformer, Copula function, principal component analysis, multi-head attention

区域经济是国民经济体系中至关重要的一部分,由于经济、政治、文化和社会的不同,区域的经济发展表现出巨大差异,导致各项经济指标表现出高度的非线性.通常情况下,对区域各个行业数据指标的预测分析,可对该区域经济发展策略的制定起到关键性的作用<sup>[1]</sup>.然而,影响区域经济的数据指标存在噪声,导致传统的预测方法很难适应区域经济预测的要求<sup>[2]</sup>.

收稿日期:2024-07-30.

基金项目:广东省哲学社会科学规划项目(GD22XYJ28)、广东省教育厅项目(2021GXJK595, 2021TSZK021).

通讯作者:刘海宏,博士,副教授,研究方向:人工智能与大数据,经济管理. E-mail:8928114@qq.com

传统的预测方法使用线性模型对区域经济进行预测. Capello 等<sup>[3]</sup>应用线性回归和高斯过程回归,利用主观因素来预测区域经济. Li 等<sup>[4]</sup>使用 SPACSYS 模型研究了经济、文化等因素对区域经济的影响. Youssef 等<sup>[5]</sup>分别利用 ARIMA 模型对未来区域的发展进行了预测,但是预测的准确率均低于 80%. Zhou 等<sup>[6]</sup>使用灰色季节模型 (grey seasonal model, GSM) 来预测区域工业的产值. 上述经济领域传统方法不能较好地拟合非线性和复杂因素的区域经济,导致预测准确率不高. 为了提高预测的准确率, Yu 等<sup>[7]</sup>通过主分量分析技术 (principal components analysis, PCA) 对影响因素进行降维,并利用支持向量机 (support vector machine, SVM) 进行区域经济预测,准确率达到 81.2%. Alaminos 等<sup>[8]</sup>将各影响指标作为随机森林的子节点,未来经济作为父节点,实现对区域经济的预测. 神经网络相较于传统的机器学习算法,能较好地处理非线性、小样本和过拟合等问题. Wang 等<sup>[9]</sup>分析了影响区域经济发展的指标,并将其输入到反向传播 (back propagation, BP) 神经网络中进行预测. 高振斌等<sup>[10]</sup>结合支持向量回归和遗传算法来预测区域的 GDP. Zhang 等<sup>[11]</sup>比较了多层感知器 (multilayer perceptron, MLP) 和长短期记忆网络 (long short-term memory, LSTM) 预测区域农业产值的性能,结果表明 LSTM 的预测性能更优.

LSTM 虽然在区域经济预测中应用广泛,但是 LSTM 的训练需要耗费巨大的计算资源,因此需要更高效的方法来对区域经济进行预测<sup>[12]</sup>. Transformer 因其高效的注意力机制为区域经济的复杂时序指标预测开辟了新思路. Laborda 等<sup>[13]</sup>利用 Transformer 的编码器-解码器结构实现对区域产值的预测. Han 等<sup>[14]</sup>结合 Transformer 和递归网络实现对区域时序指标的预测,耗费了大量的计算资源. 传统的 Transformer 模型由于局部信息获取不强、计算复杂度高,需要进一步优化才能适用于时序预测. 因此本文通过 Copula 函数对 Transformer 模型进行优化,并结合多头注意力机制对 Transformer 进行并行化处理,将优化后的 Transformer 模型应用于区域经济预测,以提高现有研究应对众多经济指标预测的准确率.

## 1 基础知识

### 1.1 Transformer 模型

Transformer 模型使用自注意力结构替代了自然语言任务中常用的循环神经网络 (recurrent neural network, RNN) 结构. 相比于 RNN 结构,其最大的优点是能灵活地捕捉长距离时序依赖关系<sup>[15]</sup>. Transformer 的模型架构如图 1 所示.

(1) 编码器. 由自注意力层 (self-attention layer) 和前馈全连接层 (feedforward layer) 组成. 自注意力层将输入序列中的每个位置与所有其他位置进行交互,前馈全连接层则将每个位置的上下文表示向量映射到另一个向量空间,以捕捉更高级别的特征.

(2) 解码器. 由掩码自注意力层、自注意力层、前馈全连接层和全连接层组成. 其中自注意力层和前馈全连接层的作用与编码器相同,而掩码自注意力层则将解码器当前位置的输入与编码器的位置进行交互,以获得与目标序列有关的信息.

### 1.2 Copula 函数

Copula 函数把边缘分布函数与联合分布函数联系起来,是研究多个变量间相依性的一种有效工具<sup>[16]</sup>. 在形式上, Copula 是一个函数值的边缘分布  $H(x)$  加上一些参数  $\vartheta$ ; 在二元情况下,它有两个边缘分布函数  $H_1(x_1)$  和  $H_2(x_2)$ . Copula 和边际分布之间的联系由 Sklar 定理<sup>[17]</sup>明确表示,对于任何联合密度,都存在连接边际密度的 Copula 如下:

$$h_{12} = h_1(x_1) \times h_2(x_2) \times c_{12}[H_1(x_1), H_2(x_2); \vartheta], \quad (1)$$

其中  $c_{12}[\cdot]$  代表变量  $x_1$  和  $x_2$  的协变量函数.

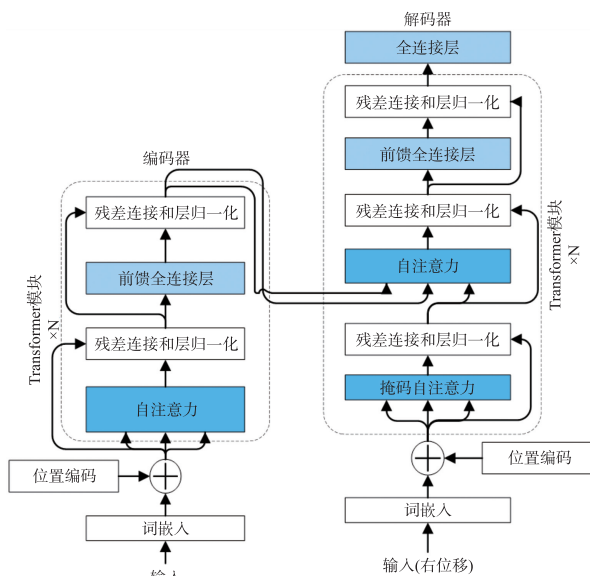


图 1 Transformer 的网络架构

Fig. 1 Network architecture for Transformer

在独立情况下, Copula 函数是常数,总是等于 1. 因此, Copula 包含了所有关于随机变量之间依赖性质的信息,但没有给出关于边际分布的信息. 例如,考虑一个具有  $N$  个随机变量的正态多元密度. 联合正态密度可以写成向量  $x$ , 均值  $\eta$  和协方差矩阵  $\Phi$  的函数.

$$h^N(x) = \frac{\exp\left[-\frac{1}{2}(x-\eta)\Phi(x-\eta)\right]}{(2\pi)^{N/2}|\Phi|^{1/2}}. \quad (2)$$

## 2 基于 Copula 函数的 Transformer 时序算法

针对 Transformer 模型缺乏可解释性以及处理稀疏时序数据效果差的问题,本文利用 Copula 函数的联合分布对其进行优化(Coformer),使得 Transformer 模型能适用于时序数据的预测. Coformer 模型主要由联合编码器、解码器组成.

(1)联合编码器. Coformer 利用 Copulas 函数中多变量边际联合分布的原理将所有标记的样本(包含缺失值)进行联合编码.

首先嵌入输入变量. 编码器为每个时序中的元素生成一个嵌入变量  $e_{ij}$ , 包含其值  $x_{ij}$ 、协变量  $c_{ij}$  以及缺失标记值  $m_{ij}$ . 该嵌入是由带有参数  $\theta$  的神经网络计算得到的,其中  $x_{ij}+m_{ij}$  隐藏了缺失标记的值.

$$e_{ij} = E_{\theta}(x_{ij}, c_{ij}, m_{ij}) = (x_{ij} + m_{ij}) \cdot c_{ij}. \quad (3)$$

然后通过位置编码  $p_{ij}$  将样本标记的时间戳  $t_{ij}$  添加到输入变量中. 为了提高编码效率,本文基于各种频率的正弦和余弦函数进行位置编码,并将最终的嵌入结果作为  $e'_{ij} = e_{ij} \sqrt{d_e} + p_{ij}$ , 其中  $d_e$  为当前位置.

(2)解码器. 由 Copula 函数的定义可知,其目的是根据观察到的标记值( $m_{ij}=1$ )、协变量( $c_{ij}$ )和时间戳( $t_{ij}$ ),学习缺失标记值( $m_{ij}=0$ )的联合分布( $x_{ij}$ ). 本文通过一个基于注意力的解码器来实现这一目标,该解码器经过训练,可以模拟非参数 copula,下面将对其进行具体描述.

数据集由能够观察到的标记和缺失的标记组成,假设  $Z^{(m)} = \{z_1^{(m)}, \dots, z_n^{(m)}\}$ , 那么  $X^{(o)} = \{x_1^{(o)}, \dots, x_n^{(o)}\}$  和  $X^{(m)} = \{x_1^{(m)}, \dots, x_n^{(m)}\}$  分别表示已观察和缺失标记的编码表示;  $C = \{c_1, \dots, c_n\}$  和  $T = \{t_1, \dots, t_n\}$  分别表示所有协变量和时间戳的集合. 解码器的目标是使用  $h_{\varphi}(x_1^{(m)}, \dots, x_n^{(m)})$  精确估计缺失标记值的联合密度.

$$h_{\varphi}(x_n^{(m)}) \approx \frac{\exp\left[-(x_n^{(o)} - \eta)\Phi(x_n^{(m)} - \eta)\right]}{(\pi)|\Phi|}, \quad (4)$$

$$\varphi = \text{concat}(X^{(o)}, Z^{(o)}, Z^{(m)}), \quad (5)$$

其中分布参数  $\varphi$  由  $X^{(o)}$ 、 $Z^{(o)}$  和  $Z^{(m)}$  拼接生成,  $\text{concat}()$  表示拼接操作.

使用编码的标记表示( $Z^{(o)}$  和  $Z^{(m)}$ )作为解码器的输入,可确保最终密度取决于实际值. 对于联合密度  $h_{\varphi}$  的 copula 计算方法如下所示.

$$h_{\varphi}(x_1^{(m)}, \dots, x_n^{(m)}) = c_{\varphi}(F_{\varphi_1}(x_1^{(m)}), \dots, F_{\varphi_n}(x_n^{(m)})) \times f_{\varphi_1}(x_1^{(m)}) \times \dots \times f_{\varphi_n}(x_n^{(m)}), \quad (6)$$

式中:  $\varphi = \{\varphi_1, \dots, \varphi_n\}$ ,  $c_{\varphi}$  代表 copula 密度,  $F_{\varphi_n}$  和  $f_{\varphi_n}$  分别代表  $x_n^{(m)}$  的边缘累积分布和密度函数.

解码器利用注意力机制为  $c_{\varphi}$  生成参数  $z$  以及累积分布函数的估计值  $\{u_1, \dots, u_n\}$ . 首先分别使用随机参数  $a$ 、 $b$  和  $c$  为每个元素( $z, u$ )生成键  $k = \text{Key}_a(z, u)$  和值  $v = \text{Value}_b(z, u)$ , 然后计算查询  $q = \text{Query}_c(z)$ . 可计算得到样本最终的注意力表示  $z'$ .

$$z' = \text{LN}(\text{FFward}(z)), \quad (7)$$

其中  $\text{LN}$  代表标准归一化层,  $\text{FFward}$  代表前馈层.

## 3 改进 Transformer 时序算法的区域经济预测

区域经济的影响指标繁多,且存在冗余,导致预测的准确率不高,针对该问题,本文提出了基于 Coformer 的区域经济预测模型,整体模型如图 2 所示. 首先选取区域经济的影响指标,对其进行预处理;其次对指标变量进行主成分分析;然后将降维后的指标变量和静态变量作为 Coformer 的输入,利用门控残差网络对各变量进行选择,并通过 Coformer 进行变量编码;最后利用解码器对编码的变量解码,通过 Softmax 输出区域历年生产总值的预测结果.

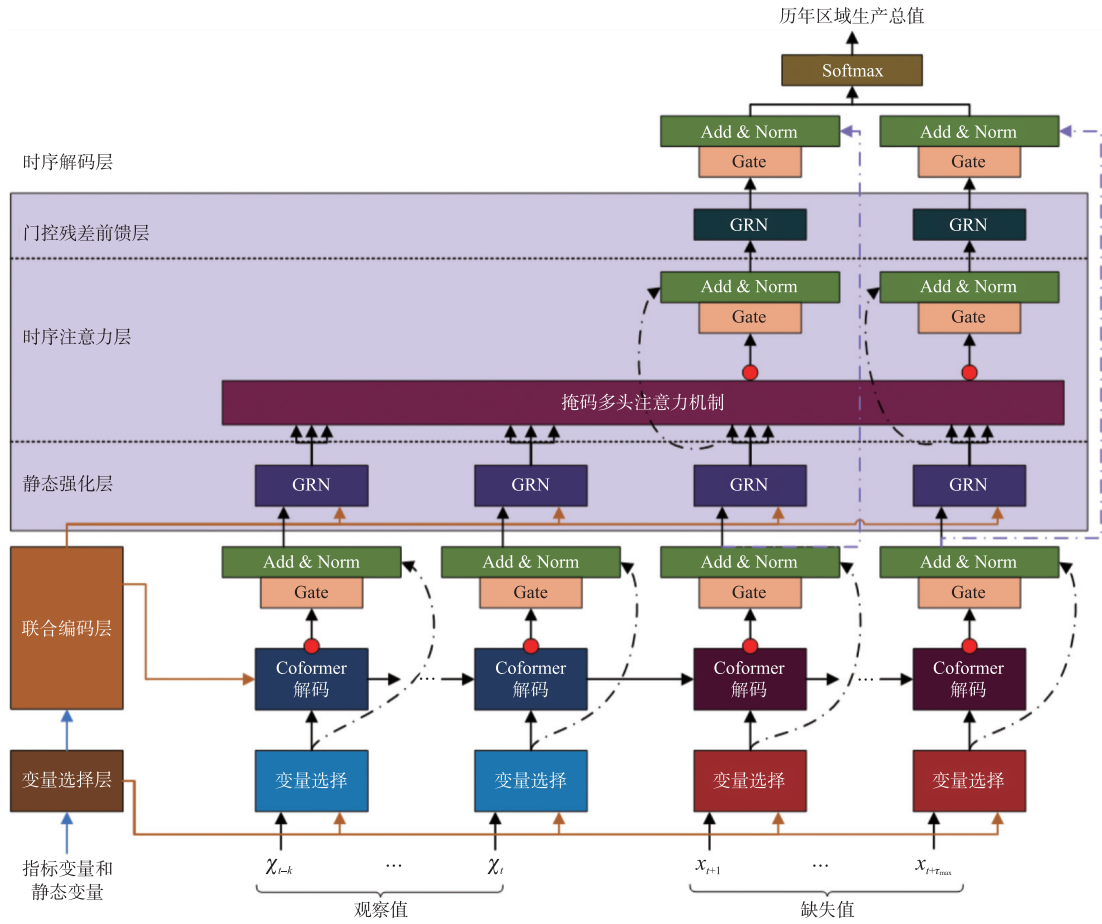


图 2 基于 Coformer 的区域经济预测模型结构

Fig. 2 Regional economic forecasting model based on Coformer

### 3.1 区域经济影响指标的选取及降维

为了实现对区域经济的准确预测,需要对影响区域经济的指标进行分析. 根据现有研究<sup>[18]</sup>可知,经济指标的选取总共包括 6 大类:生产指标类、物价指标类、财政指标类、消费指标类、贸易指标类和资源指标类,17 个能够反映区域经济发展态势的数据指标:第一、二和三产业生产总值、全社会零售物价指数、政府财政支出、政府财政收入、全社会零售品销售额、零售物价指数、固定资产投资、实际利用外资、外贸进口额、外贸出口额、金融机构存款余额、能源消耗总量、教育投入、环保投入、科技投入,分别记为  $x_1, x_2, \dots, x_{17}$ ,将这些指标作为预测模型的输入变量,将区域历年生产总值  $y$  作为输出变量.

为了消除输入变量的量纲对预测结果的影响,本文对选取的指标进行归一化处理,使得各变量都压缩在  $[0, 1]$  的范围内.

$$\bar{x}_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}. \quad (8)$$

为了去除影响指标中的冗余信息,利用 PCA 选择贡献率较大的主成分作为预测模型的输入变量,具体步骤如下.

Step 1:对各个指标进行标准化处理,即  $y_{ij} = \frac{(x_{ij} - \bar{x}_j)}{r_j}$ ,其中  $\bar{x}_j$  和  $r_j$  分别表示不同维度分量的均值和标准差.

Step 2:基于标准化后的数据,计算变量之间的协方差矩阵  $R$ .

$$R = \frac{[Y - \bar{Y}][Y - \bar{Y}]^T}{N}, \quad (9)$$

其中,  $\bar{Y} = [\bar{y}_1, \bar{y}_2, \dots, \bar{y}_M]$ .



Step 3: 对特征方程  $(\lambda I - R)V = 0$  进行特征值分解, 其中  $l = [1, 1, \dots, 1]_{1 \times N}$ , 求得特征值  $\mu$  和相应的特征向量  $V$ .

Step 4: 将特征值依据从大到小的顺序进行排列, 选择前  $l$  个特征向量作为主成分  $X_j$ , 使用选定的主成分构建新的特征空间. 每个数据点都可以通过将其投影到这些主成分上来表示, 从而实现数据的降维.

### 3.2 区域经济预测过程

在对影响指标进行降维处理后, 本文基于 Coformer 时序算法实现对区域经济的预测, 该预测模型包含变量选择层、联合编码层、多头注意力层、时序解码层。

(1) 变量选择层. 为了提高模型的灵活性, 需对影响指标采取非线性化处理, 利用门控残差网络 (gated residual network, GRN) 中的门控线性单元 (gated linear unit, GLU) 代替 Transformer 中的 ReLU 激活函数. GRN 将降维后的变量  $X_j$  和静态变量  $a_i$  作为输入, 通过 GRN 中的门控组件过滤无关变量, 如式 (10) 所示. 这里的静态变量包括工业、农业、服务业等行业变量和所在区域、城区等位置变量, 这些变量与时间不相关.

$$GRN_{\omega}(X_j, a_i) = \tanh[(X_j w + a_i) \otimes a_i], \quad (10)$$

其中  $\tanh$  是线性单元激活函数,  $\omega$  代表不同层之间共享的权重,  $\otimes$  表示克罗内克积.

使用 GLU 的门控层抑制模型中不关注的变量, 以静态变量  $a_i$  为例, 过滤结果如下:

$$GLU_{\omega}(a_i) = \delta(w a_i + b) \odot a_i, \quad (11)$$

其中  $\delta(\cdot)$  代表 sigmoid 激活函数,  $b$  代表偏置,  $\odot$  代表哈达玛乘积.

(2) 联合编码层. 在获得最终的指标变量和静态变量后, 需对其特征进行提取. 通过联合编码器捕捉静态变量和影响指标变量的时序特征, 将时序特征输入到 Coformer 进行编码, 以学习各个变量间的相关性, 提高区域经济预测的准确性. 在现有研究中, 数据集都缺失了静态变量, 利用第二章的 Coformer 对指标变量和静态变量进行联合分布编码, 得到编码结果为  $e_{ij}^*$ .

为了增强各变量中的关键信息, 降低计算开销, 使用多头注意力机制提高重要变量对区域历年生产总值的影响, 利用多头注意力机制, 对静态变量和指标变量采用不同的头, 如下所示:

$$MHead(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) V W, \quad (12)$$

其中  $MHead(\cdot)$  代表多头注意力机制,  $Q, K, V$  分别代表查询、键和值,  $d_k$  表示  $Q, K, V$  的维度,  $W$  代表多头注意力机制中头部的权重值.

鉴于每个头部使用不同的值, 仅注意力权重并不能指示特定特征的重要性. 因此, 所提注意力机制在每个头中共享值, 并对所有头进行叠加聚合, 得到最终的输出值  $\tilde{H}$ .

$$\tilde{H} = \frac{\sum_{i=1}^m \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right) V W_i}{m}, \quad (13)$$

其中  $W_i$  代表第  $i$  个头部的权重,  $m$  表示头部的数量.

随后, 利用 GRN 对多头注意力的输出  $\tilde{H}$  进行非线性化处理, 编码得到最终的输出  $O_E$  如下.

$$O_E = LN(GRN_{\omega}(\tilde{H})) + \tilde{H}, \quad (14)$$

其中  $LN$  代表标准归一化层,  $GRN_{\omega}(\tilde{H}) = \tanh(\tilde{H} w)$ .

(3) 时序解码层. 在对输入变量进行编码后, 解码器需对编码的变量解码, 获得区域历年生产总值的预测结果. 时序解码层包括静态强化层、时序注意力层、门控残差前馈层.

第一层为静态强化层, 由于静态变量通常会对时序特征产生重大影响, 因此我们引入了一个静态丰富层, 利用静态变量增强时间特征. 对于给定的位置索引  $n$ , 静态变量强化的结果如下.

$$\theta(t, n) = GRN_w(\varphi(t, n), a_e) = \tanh[(\varphi(t, n) w + a_e) \otimes a_e], \quad (15)$$

其中  $a_e$  是静态变量的编码.

第二层为时序注意力层. 对所有静态变量强化的时间特征进行组合, 即  $\Theta(t) = [\theta(t, 1), \dots, \theta(t, n)]^T$ , 然后在每个预测时间内应用多头注意力机制  $\beta(t)$ , 从而得到解码器解码的时序信息  $B'(t)$ .

$$\beta(t) = MHead(\Theta(t), \dots, \Theta(t)), \quad (16)$$

$$B'(t) = [\beta(t, 1), \dots, \beta(t, n)]. \quad (17)$$

注意力层采用了解码器掩码技术,以确保每个时间维度只能关注关键特征,并且能让模型捕捉到输入变量和预测产值之间的依赖关系。

第三层为门控残差前馈层. 对时序注意力层的输出进行非线性处理. 利用门控残差连接各个时序序列,无需额外的计算消耗。

$$\eta = LN(B'(t) + GLU_w(B'(t), a_i)), \quad (18)$$

其中  $GLU_w(B'(t), a_i) = \delta(B'(t)) \odot a_i$ .

通过 GRN 组件获得解码器的输出  $O_D$ .

$$O_D = LN(\eta + GRN_w(\eta)), \quad (19)$$

其中  $GRN_w(\eta) = \tanh(\eta w)$ .

最后,对  $O_D$  进行线性变换,通过 Softmax 输出获得预测的区域历年生产总值序列  $O_S$ .

$$O_S = \text{Softmax}(\text{Linear}(O_D)), \quad (20)$$

其中  $\text{Softmax}(\cdot)$  是概率归一化激活函数,  $\text{Linear}(\cdot)$  是线性映射函数。

## 4 实验与结果分析

### 4.1 区域生产总值预测结果分析

为了评估所提模型在区域经济中的预测效率,选取某市统计年鉴中 4 个区 2010 到 2023 年的经济数据及影响因素. 数据集中的因变量为区域历年生产总值,自变量为 PCA 所选出的主要指标. 将统计结果整理好后,利用 MATLAB R2016a 进行实验统计分析. 本文将经济数据样本集按照 7:2:1 的比例划分训练集、测试集和验证集,对本文设计的预测模型 Coformer 和 REDBP 模型<sup>[9]</sup>、SVRGA 模型<sup>[10]</sup>、Eformer 模型<sup>[14]</sup>进行对比和分析. 所有模型的训练精度为 0.001,学习率为 0.001,最大训练次数为 100 次。

分别利用 Coformer、REDBP、SVRGA 和 Eformer 对 2010—2020 年某市四个区域的生产总值进行训练学习,实验结果如图 3 所示. 由下图可知,采用 Coformer 对 2010—2020 年的区域生产总值进行预测,其预测值与区域实际生产总值之间的误差最小. 而 REDBP 和 SVRGA 的预测结果较差,这是因为 REDBP 仅利用 BP 神经网络模型对区域生产总值进行预测,忽略了指标变量和静态变量的内在关系. SVRGA 采用 ANN 预测区域生产总值,未对影响指标去除冗余信息,导致预测结果出现偏差. Eformer 是基于 Transformer 算法和递归网络进行预测,但是未对传统 Transformer 算法优化,也未对这些变量进行降维处理,使得其预测结果差于 Coformer. 因此,Coformer 相比于另外三种模型表现出了优异的预测性能。

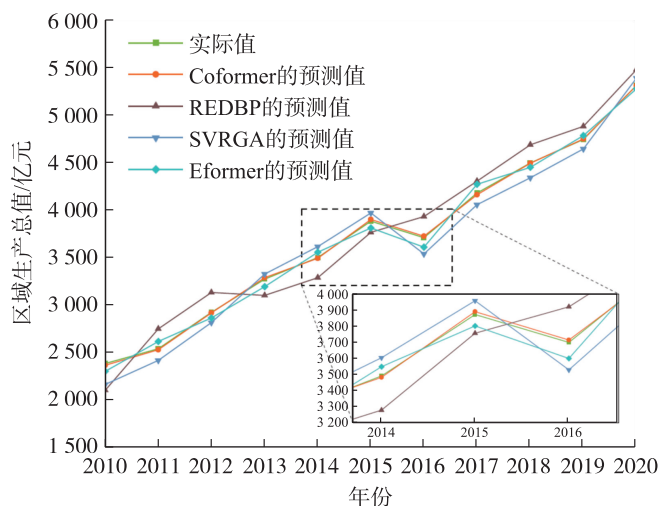


图 3 不同模型的区域生产总值预测结果

Fig. 3 Regional gross domestic product forecast results of different models

利用 2010—2020 年数据集训练获得的 Coformer、REDBP、SVRGA 和 Eformer 对 2021—2023 年的测试集进行区域生产总值预测,实验结果如表 1 和表 2 所示. 对区域生产总值的实际值的预测,Coformer 和 Eformer 在 2023 年绝对误差最小,在 2022 年的绝对误差最大. REDBP 和 SVRGA 在 2023 年绝对误差最小,在 2021 年的绝对误差最大. 从 2021—2023 年区域生产总值的误差率来看,Coformer 的预测误差率分别为 1.1%、1.3%和 0.8%,相比于 REDBP 分别降低了 3.9%、2.9%和 2.2%,相比于 SVRGA 分别降低了 3.0%、1.6%和 1.8%,相比于 Eformer 分别降低了 0.7%、0.6%和 0.6%,充分验证了所提模型 Coformer 的高效性. Coformer 不仅通过 PCA 去除了输入变量的冗余数据,而且利用 Copula 函数对 Transformer 进行优化,使得优化后的 Coformer 更加关注重要的时序序列,从而进一步提高了预测效率.

表 1 四种模型对区域生产总值的预测结果

Table 1 Forecasting results of the four models for gross regional product

年度	实际值/亿元	预测值/亿元			
		Coformer	REDBP	SVRGA	Eformer
2021	5 497.62	5 438.16	5 771.37	5 725.28	5 397.93
2022	5 708.37	5 632.84	5 997.60	5 544.12	5 818.21
2023	6 125.91	6 072.30	6 307.93	5 969.44	6 211.11

表 2 四种模型的预测误差

Table 2 Prediction errors of the four models

年度	绝对误差/亿元				误差率/%			
	Coformer	REDBP	SVRGA	Eformer	Coformer	REDBP	SVRGA	Eformer
2021	59.46	273.75	227.66	99.69	1.1	5.0	4.1	1.8
2022	75.53	239.23	164.25	109.84	1.3	4.2	2.9	1.9
2023	53.61	182.02	156.47	85.2	0.8	3.0	2.6	1.4

4.2 预测性能分析

除了通过绝对误差和误差率直观对比各个模型的预测效率外,也可通过预测性能评价指标:准确率 (ACC)、召回率 (REC)、精准率 (PREC)、F1 (REC 和 PREC 的调和平均值) 和 ROC 曲线下方的面积 (AUC)对模型进行深入分析,从而得到更准确的评估结果. 各模型的预测性能指标对比如图 4 所示.

由图 4 可知,Coformer、REDBP、SVRGA 和 Eformer 的准确率分别为 0.908、0.749、0.785 和 0.841,均大于 0.7,表明 4 种模型的预测性能都挺高. Coformer 的准确率最高,相比于 REDBP、SVRGA 和 Eformer 分别提高了 15.9%、12.3%和 6.7%. 仅比较准确率具有片面性,F1 是召回率和精准率的综合值,使得评价更具全面性. Coformer 的预测能力最强,F1 达到了 0.905,SVRGA 和 Eformer 的 F1 值都处在 0.8 至 0.9 之间,REDBP 的 F1 最低,仅为 0.75.

AUC 值是评价预测性能的重要指标,AUC 值越大,预测效果越佳. Coformer 的 AUC 值为 0.934,最接近 1,预测性能最好. REDBP、SVRGA 和 Eformer 的预测性能虽然弱于 Coformer,但是 AUC 值均大于 0.5,表明各模型的有效性和 Coformer 的优越性.

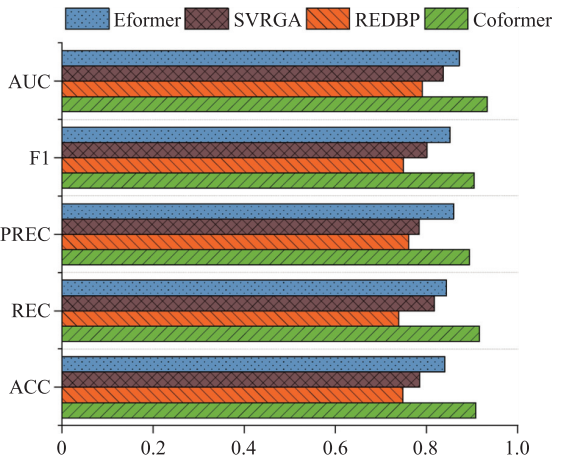


图 4 不同模型的预测性能指标对比

Fig. 4 Comparison of predictive performance indexes of different models

5 结论

区域经济预测是一项极具挑战性的任务,本文针对现有区域经济预测方法准确率低下的问题,设计了基于改进 Transformer 时序算法的区域经济预测模型. 该模型具有以下优势:

- (1)采用 Copula 函数对传统 Transformer 模型进行优化,使其能适用于时序数据的预测.

(2)对选取的区域经济影响指标进行 PCA 降维,去除冗余信息.

(3)利用 Coformer 对指标变量和静态变量进行编码和解码,基于门控残差网络关联输入变量和预测序列;结合多头注意力机制增强重要信息,通过 Softmax 输出区域历年生产总值的预测结果.

(4)所提方法的预测准确率和 AUC 值分别为 0.908、0.934,预测效率更高.

## [参考文献]

- [1] 邹子建. 基于 BP 神经网络的衡水市经济预测研究[D]. 咸阳:西北农林科技大学,2010.
- [2] LEHMANN R,WOHLRABE K. Boosting and regional economic forecasting;the case of Germany[J]. Letters in spatial and resource sciences,2017,10:161-175.
- [3] CAPELLO R. A forecasting territorial model of regional growth;the MASST model[J]. The annals of regional science,2007,41:753-787.
- [4] LI M,SUN H,AGYEMAN F O,et al. Analysis of potential factors influencing China's regional sustainable economic growth[J]. Applied sciences,2021,11(22):10832.
- [5] YOUSSEF J,ISHKER N,FAKHREDDINE N. Economic forecast of the wealthiest gulf countries using ARIMA model[J]. Journal of applied economic sciences,2021,16(2):228-237.
- [6] ZHOU W,CHENG Y,DING S,et al. A grey seasonal least square support vector regression model for time series forecasting[J]. ISA transactions,2021,114:82-98.
- [7] YU N,XU W,YU K L. Research on regional logistics demand forecast based on improved support vector machine;a case study of Qingdao city under the New Free Trade Zone Strategy[J]. IEEE access,2020,8:9551-9564.
- [8] ALAMINOS D,BECERRA-VICARIO R,FERNÁNDEZ-GÁMEZ M á,et al. Currency crises prediction using deep neural decision trees[J]. Applied sciences,2019,9(23):5227.
- [9] WANG J,LI Y,SONG J. Research on regional economic development difference forecast based on BP neural network[J]. Forest chemicals review,2022:952-963.
- [10] 高振斌,秦轩. 基于 PLS-GA-SVR 的区域经济预测模型及实证[J]. 统计与决策,2022,38(18):54-57.
- [11] ZHANG K,THÉ J,XIE G,et al. Multi-step ahead forecasting of regional air quality using spatial-temporal deep neural networks;a case study of Huaihai Economic Zone[J]. Journal of cleaner production,2020,277:123231.
- [12] WAN Y,TAO H,MA L. Forecasting Zhejiang Province's GDP using a CNN-LSTM model[J]. Frontiers in business, economics and management,2024,13(3):233-235.
- [13] LABORDA J,RUANO S,ZAMANILLO I. Multi-country and multi-horizon GDP forecasting using temporal fusion transformers[J]. Mathematics,2023,11(12):2625.
- [14] HAN Y,TIAN Y,YU L,et al. Economic system forecasting based on temporal fusion transformers; multi-dimensional evaluation and cross-model comparative analysis[J]. Neurocomputing,2023,552:126500.
- [15] 周湘贞,李帅,隋栋.基于深度学习和注意力机制的微博情感分析[J]. 南京师大学报(自然科学版),2023,46(2):115-121.
- [16] 宋继勳,周春雷,沈子奇,等. 基于 TCN-GRU 的 Handle 标识解析系统负载均衡算法[J]. 福建师范大学学报(自然科学版),2024,40(2):64-73.
- [17] 李庆涛,林培光,王基厚,等. 基于板块效应的深度学习股价走势预测方法[J]. 南京师范大学学报(工程技术版),2022,22(1):30-38.
- [18] 项晓宇,朱敏捷,周灵刚,等. 基于机器学习的短期规上行业工业增加值预测[J]. 南京师大学报(自然科学版),2023,46(2):99-106.

[责任编辑:杜忆忱]