

# 时间序列预测：方法、进展与展望

卢斌龙

福建师范大学计算机与网络空间安全学院，福建福州，350117；

**摘要：**时间序列数据因其内在的时间依赖性，在广泛领域具有关键价值。本文阐述了时间序列预测的理论框架，追踪其方法论从传统统计学到机器学习、深度学习的演进路径。深度学习模型，如CNN、RNN和Transformer，在捕捉数据复杂模式和预测准确性方面表现卓越。文章系统性梳理了时间序列数据特性和预测任务类型，并展望了异常检测、超参数优化等未来研究方向，以推动深度学习在时间序列分析中的创新应用。

**关键词：**时间序列预测；自回归模型；深度学习

DOI:10.69979/3041-0673.24.6.037

## 引言

时间序列是按时间顺序排列的一组随机变量，通常以固定时间间隔对某一过程进行观测。广泛存在的时间序列数据涵盖金融市场、气象监测、等领域，在科学研究中具有重要意义。近年来，传感器和网络技术的进步使得大量时间序列数据得以生成和积累。

分类聚类、异常检测、事件预测、时间序列预测是时间序列数据的四个重点研究方向。本文聚焦于时间序列预测。与传统回归模型不同，时间序列预测依赖于数据的时间序列关系，目标是通过分析历史数据规律，预测未来趋势。其研究内容包括连续型和离散型预测，具有重要价值。

随着数据规模和复杂性的增加，时间序列预测方法经历了从数学统计到机器学习的演进。尤其是深度学习领域。本文首先对于时间序列预测问题进行了讨论，然后对于其研究方法进行了梳理，在最后对于时间序列预测技术的发展进行了进一步展望

## 1 时间序列预测概述

### 1.1 时间序列数据特性

时间序列数据因其独特的时间依赖性而区别于其他类型的数据。以下是时间序列数据的几个核心特性，这些特性对于构建和评估时间序列预测模型至关重要。

1. 海量性：时间序列数据的规模随着技术进步而急剧增加。这种数据的爆炸性增长要求我们在预处理阶段进行有效的数据管理和处理。

2. 趋势性：时间序列数据常表现出长期的发展态势。趋势性揭示了数据随时间变化的总体方向，这种变化可

能是线性或非线性的，对预测模型的长期预测能力提出了要求<sup>①</sup>。

3. 周期性：许多时间序列数据表现出周期性波动。周期性特征对于理解数据的规律性变化至关重要，对于预测模型捕捉这些规律性模式具有指导意义。

4. 季节性：季节性是周期性的一种特殊形式，它描述了数据在固定周期内重复出现的模式。例如，零售业的销售数据常在节假日前后出现高峰，这种季节性变化对于预测模型的准确性至关重要。

5. 平稳性：

平稳性描述了时间序列数据的统计属性在时间上是否保持恒定。非平稳时间序列可能需要通过差分、对数转换等方法进行转换，以便于分析和建模。

在构建时间序列预测模型时，深入理解这些特性对于捕捉数据的内在规律和预测未来变化至关重要。通过综合考虑这些特性，可以设计出更加精确和可靠的预测模型。

### 1.2 时间预测预测任务分类

在时间序列预测领域，根据不同的特征和需求，预测任务可以分为几个主要类别。以下是对这些类别的描述和它们各自的难点：

#### 1.2.1 单变量与多变量时间序列预测

单变量时间序列预测指基于单一时间序列变量的数据对未来趋势进行预测。例如，利用历史股价预测未来价格。其核心在于分析时间序列的内部特性，如趋势、季节性和平稳性。趋势反映长期运动方向，季节性描述定期出现的周期性波动，平稳性则指序列的统计特性不随时间变化。因此，噪声与非平稳性的存在增加了捕捉

这些特性及构建精确模型的难度。多变量时间序列预测则利用多个相关时间序列变量,通过分析变量间的动态交互来预测未来值。这种方法能够捕捉变量间复杂的依赖关系,从而提升预测性能。但高维数据的处理、多变量之间的相互作用,以及异质数据的整合等问题为其带来了巨大挑战。

### 1.2.2 短期与长期时间序列预测

短期时间序列预测关注于预测未来几个时间步的值,而长期时间序列预测则关注于更远未来的值。短期预测通常更容易实现,因其受到的影响因素较少,且这些因素更容易被观测和建模。长期预测则更具挑战性,因为它需要考虑更多的不确定性和变化,以及如何处理时间序列数据中的非平稳性和结构变化,此外,长期预测容易受到外部因素(如政策变化)的影响,这些因素难以在模型中准确量化和预测。

### 1.2.3 平稳与非平稳时间序列预测

平稳时间序列预测假设时间序列的均值和方差在时间上保持恒定,且自相关结构不随时间变化,这使得建模和预测变得更为简单和直接,但在实际应用中,实际数据仍可能受到随机波动和噪声的影响。此外,模型选择和参数估计的准确性也对预测效果有重要影响。非平稳时间序列预测则需要处理时间序列的统计特性随时间变化的情况,其预测的难点在于如何有效地识别和处理非平稳性,以及如何在模型中捕捉时间序列的动态变化。此外,非平稳时间序列中的突发事件和异常值也会对预测模型产生较大影响。

### 1.2.4 稳健性时间序列预测

稳健性时间序列预测关注的是构建能够在各种不确定性和噪声条件下保持稳定性和准确性的预测模型。这种类型的预测任务需要模型能够抵抗异常值、噪声和模型不确定性的影响,从而在数据存在污染或异常时,仍然提供可靠的预测结果。稳健性时间序列预测的难点在于如何设计和优化模型以提高其对异常值和噪声的鲁棒性,以及如何在模型中合理地刻画和应对不确定性。

## 2 时间序列预测方法

### 2.1 传统时间序列预测方法

传统时间序列预测方法主要是在确定时间序列参数模型的基础上,求解出模型参数,并利用求解出的模型完成未来预测工作。这些方法主要基于统计学原理,

通过分析历史数据来预测未来的数据点。以下是几种常用的传统时间序列预测方法:

#### 1. ARIMA (自回归积分滑动平均模型)

ARIMA 模型是时间序列预测中最为广泛使用的方法之一。它结合了自回归 (AR)、差分 (I) 和移动平均 (MA) 三个部分,适用于非平稳时间序列数据的预测。ARIMA 模型通过拟合历史数据来识别数据中的模式,并据此预测未来的值。参数模型的阶数  $(p, d, q)$  是 ARIMA 模型的关键,其中  $p$  代表自回归项的阶数,  $d$  代表差分阶数,  $q$  代表移动平均项的阶数。

#### 2. Holt-Winters 法

Holt-Winters 法,也称为三指数平滑法,是专门用于具有趋势和季节性的时间序列数据的预测方法。它通过三个平滑系数  $(\alpha, \beta, \gamma)$  分别对水平、趋势和季节性成分进行建模,以适应数据的变化。Holt-Winters 法在处理具有明显趋势和季节性的数据时表现良好。

#### 3. Box-Jenkins 方法

Box-Jenkins 方法是一种系统化的时间序列建模方法,包括模型识别、参数估计、模型检验和预测四个阶段。该方法首先需要将非平稳序列转化为平稳序列,并确定合适的 ARMA 模型阶数,然后估计模型参数,进行模型诊断,最后利用模型进行预测。

#### 4. 状态空间方程

状态空间模型是一种线性动态系统,它将时间序列数据建模为一个隐藏状态过程和一个观测过程。卡尔曼滤波是状态空间模型中的关键技术,用于递归地估计隐藏状态,并预测未来的观测值。

尽管这些传统方法在某些场景下非常有效,但它们也存在一些局限性。例如,对非线性关系和复杂结构的捕捉能力有限,且对异常值和噪声较为敏感等。因此,随着深度学习技术的发展,基于深度学习的时间序列预测方法受到越来越多的关注。

## 2.2 基于深度学习的时间序列预测方法

深度学习技术在时间序列预测领域相较于传统算法展现出了卓越的性能,其强大的特征提取能力使其能够捕捉到传统方法可能忽视的模式和规律,从而满足高精度预测的需求。

### 2.2.1 基于卷积神经网络的方法

卷积神经网络 (CNN) 在图像处理领域取得了巨大成功,其在时间序列分析中同样有效,尤其是在提取局

部时间特征方面。然而，传统 CNN 在时间序列预测中的性能受限于其有限的感受野和对长距离依赖的捕捉能力。为了解决这一问题，DeepMind 提出的 WaveNet<sup>[19]</sup>模型通过空洞卷积扩展了模型的感受野，使得模型能够聚合来自更远时间点的信息。时间卷积网络 (TCN) 则通过因果卷积和空洞卷积，以及残差连接，提供了宽广的时域感受野，同时保持了并行计算的优势，提高了运算效率。SCINet 通过层次化结构在不同时间分辨率下提取特征，学习短期和长期依赖关系，TimesNet 则在周期性的基础上将时间序列重塑为二维形式，采用二维卷积的方式以捕获时间段内部依赖和相邻时间段内的依赖。

### 2.2.2 基于循环神经网络的方法

循环神经网络 (RNN) 在序列建模中具有天然优势，能够捕捉时间序列内部的隐藏状态作为过去信息的特征表征。长短期记忆网络 (LSTM) 和门控循环单元 (GRU) 作为 RNN 的变体，通过门控制机制解决了梯度消失和爆炸问题，成为时间序列预测中的关键技术。DeepAR 网络利用 LSTM 模型进行时间序列预测，通过在训练和预测阶段使用不同的输入策略，提高了预测的准确性。C2FAR 在 DeepAR 网络的基础上，巧妙的从序列值离散化的角度来解决预测问题。而 WITRAN 通过双粒度信息传输捕获长期和短期重复模式，利用水平垂直门控选择单元递归融合和选择信息，从而建立起了全局和局部关联的模型。

### 2.2.3 基于 Transformer 的方法

Transformer 模型，因其自注意力机制在处理序列信息时的高效性而被应用于时间序列预测。与传统 RNN 相比，Transformer 能够更好地处理长距离依赖关系，并通过并行计算提高训练效率。然而，Transformer 在计算和内存需求上较高。为了解决这些问题，Informer 提出了稀疏概率自注意力机制进行选择查询，自注意力蒸馏以减少维度和网格的参数数量，此外其直接生成长期预测序列以减少累积误差。Autoformer 则将序列分解引入自注意力机制，并基于基于随机过程理论，提出自相关机制以替代点向连接的注意力。PatchTST 将计算机视觉中的 patch 概念引入时序预测领域，取得了优异的成果。

### 2.2.4 基于图神经网络的方法

GNN 通过图结构学习层定义信息传播的图结构，并在其上进行卷积操作，以捕获时间序列内的空间和时间

依赖性。Graph WaveNet 通过引入自适应图建模方法和空洞因果一维卷积模型，精确地捕获数据中隐藏的空间依赖性，以提取长期时序关系。MTGNN<sup>[20]</sup>提出了自适应图学习层以从数据中抽取稀疏图邻接矩阵，并采用混合跳跃传播层和膨胀卷积层来捕捉不同时间尺度上的特征。FourierGNN 则面向以往时空分开建模的问题，通过堆叠其提出的傅立叶图算子以在傅立叶空间中执行矩阵乘法，实现了性能，效率，参数三项的提高。

## 3 展望

随着深度学习技术在时间序列预测领域的不断进步，未来的研究和应用将面临新的挑战 and 机遇。以下是几个关键的研究方向：

### 1. 超参数优化

深度学习模型的性能往往受到超参数选择的显著影响。随机自然启发优化算法，如遗传算法、粒子群优化等，可以用于自动寻找最优超参数组合，以提升模型的预测能力。这些算法模仿自然现象和生物行为，为解决深度学习中的优化问题提供了新的思路。

### 2. 小数据集与不规则时间间隔的网络架构

针对小数据集和时间间隔不规则的时间序列数据，研究者需要开发新的网络架构，以提高模型的泛化能力和适应性。这可能包括引入重采样、插值、滤波等技术，以及探索适合这类数据的轻量级模型，以解决过拟合问题并提高预测性能。

### 3. 领域知识整合与因果性推断

在不同的应用领域中，深度学习模型需要整合领域知识以提高预测的精度和效率。同时，未来的研究应关注因果性推断，提高预测结果的可解释性，而不仅仅是依赖于序列内部的相关性分析。

### 4. 模型轻量化与多模态数据融合

为了使模型在资源受限的环境中部署，模型轻量化设计成为必要。同时，多模态数据融合可以提供更全面的预测信息，提高模型的预测能力和泛化能力。

## 4 结束语

时间序列预测旨在通过分析历史数据来预测未来趋势，作为一项重要的任务。其随着应用场景和技术在不断的蓬勃发展。本文首先探讨了时间序列数据的独特特性，随后对时间序列预测任务进行了分类，介绍了传统的统计方法与新兴的基于深度学习的技术。最后，

文章在此基础上对未来的研究方向进行了展望,包括异常点检测、超参数优化和模型轻量化设计等

### 参考文献

- [1]WEN Q, ZHOU T, ZHANG C, 等. Transformers in Time Series: A Survey[A/OL]. arXiv, 2023[2024-11-22].
- [2]LIU Y, WU H, WANG J, 等. Non-stationary Transformers: Exploring the Stationarity in Time Series Forecasting[A/OL]. arXiv, 2023[2024-11-22].
- [3]OORD A van den, DIELEMAN S, ZEN H, 等. Wave Net: A Generative Model for Raw Audio[A/OL]. arXiv, 2016[2024-11-22].
- [4]BAI S, KOLTER J Z, KOLTUN V. An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling[A/OL]. arXiv, 2018[2024-11-22].
- [5]LIU M, ZENG A, CHEN M, 等. SCINet: Time Series Modeling and Forecasting with Sample Convolution and Interaction[A/OL]. arXiv, 2022[2024-11-22].
- [6]WU H, HU T, LIU Y, 等. TimesNet: Temporal 2D-Variation Modeling for General Time Series Analysis[A/OL]. arXiv, 2023[2024-11-22].
- [7]SALINAS D, FLUNKERT V, GASTHAUS J. DeepAR: Probabilistic Forecasting with Autoregressive Recurrent Networks[A/OL]. arXiv, 2019[2024-11-22].
- [8]BERGSMA S, ZEYL T, ANARAKI J R, 等. C2FAR: Coarse-to-Fine Autoregressive Networks for Precise Probabilistic Forecasting[A/OL]. arXiv, 2023[2024-11-22].
- [9]WITRAN: Water-wave Information Transmission and Recurrent Acceleration Network for Long-range Time Series Forecasting. [2024-11-22]
- [10]ZHOU H, ZHANG S, PENG J, 等. Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting[A/OL]. arXiv, 2021[2024-11-22].
- [11]WU H, XU J, WANG J, 等. Autoformer: Decomposition Transformers with Auto-Correlation for Long-Term Series Forecasting[A/OL]. arXiv, 2022[2024-11-22].
- [12]NIE Y, NGUYEN N H, SINTHONG P, 等. A Time Series is Worth 64 Words: Long-term Forecasting with Transformers[A/OL]. arXiv, 2023[2024-11-22].
- [13]WU Z, PAN S, LONG G, 等. Connecting the Dots: Multivariate Time Series Forecasting with Graph Neural Networks[A/OL]. arXiv, 2020[2024-11-22].
- [14]YI K, ZHANG Q, FAN W, 等. FourierGNN: Rethinking Multivariate Time Series Forecasting from a Pure Graph Perspective[A/OL]. arXiv, 2023[2024-11-22].