迈向准确且可解释的时间序列预测：多项式学习方法

Bo Liu[a, b]，Shao-Bo Lin[a]，Changmiao Wang[c]，Xiaotong Liu[a, *]

[a]智能决策与机器学习中心、管理学院、西安交通大学，中国西安[b]中国电子科技集团公司第三十九研究所，中国西安[c]深圳大数据研究院，中国深圳

arXiv:2603.02906v1 [cs.LG] 3 Mar 2026

摘要

时间序列预测能够实现预警，并推动资产绩效管理从传统的计划性维护转向预测性维护。然而，预测方法缺乏可解释性，削弱了用户的信任，也让开发人员更难调试。因此，可解释的时间序列预测日益受到研究关注。尽管如此，现有方法仍存在若干局限，包括对时间依赖关系的建模不足、缺乏支持预警的特征级可解释性，以及难以兼顾准确率与可解释性。本文提出可解释多项式学习（IPL）方法，通过多项式表示显式建模原始特征及其任意阶交互，将可解释性融入模型结构。该设计保留了时间依赖关系，提供特征级可解释性，并可通过调整多项式次数，在预测准确率与可解释性之间灵活权衡。我们在模拟数据和比特币价格数据上评估了 IPL，结果表明，与广泛使用的可解释性方法相比，它在实现高预测精度的同时具有更出色的可解释性。对实地采集的天线数据进行的实验进一步表明，IPL 能够实现更简单、高效的预警机制。关键词：时间序列预测；可解释性；预警；多项式学习

*通讯作者：ariesoomoon@gmail.com

1. 引言

时间序列数据按顺序记录系统状态的演变，涵盖金融市场中的股价波动、工业设备的传感器读数，以及医疗保健领域的心电监测 [1, 2, 3]。利用时间序列数据中固有的时间依赖关系，可以通过识别历史上反复出现的模式来推断未来状态；因此，时间序列预测广泛应用于医疗早期预警系统、工程预测性维护和金融市场监测等关键领域 [4, 5]。事实上，在资产绩效管理（APM）领域，重点已从传统的计划性维护转向基于时间序列预测的预测性维护[1]，；在许多行业，准确预警重大设备故障的价值超过 100，000 美元[2]。2024年，全球预测性维护市场价值为 109.3 亿美元，预计到 2032 年将增长至 707.3 亿美元，预测期内的复合年增长率为 26.5%[3]。

现有的基于时间序列的早期预警系统（如 OpenText、Falkonry、DolphinDB 和 Artesis）主要依赖两种范式：（1）由人类经验驱动的基于规则的告警；（2）根据关键指标偏离正常状态的程度进行异常评分，并据此触发警报。这两种范式都并非真正基于数据驱动的预测模型，而且都难以确保预警确实对应真实异常。因此，据 IoT Analytics 统计，目前早期预警系统的准确率大多仍低于 50%，这进一步导致告警疲劳，并削弱了操作人员的信任。与此同时，现有的时间序列预测方法——尤其是基于深度学习的方法，例如循环神经网络、卷积神经网络和 Transformer [6, 7, 8, 9, 10]——已取得最先进的预测性能；然而，其黑箱特性使得预测生成过程难以解释。除了出色的预测

[1]https://www.gelonghui.com/p/916603[2] https://iot-analytics.com/product/predictive-maintenance-asset-performance-market-report-2023-2028/[3]https://www.fortunebusinessinsights.com/press-release/predictive-maintenance-market-9602

性能之外，最终使这些模型能够在发出警告后支持可信、可执行且有针对性的干预措施的，是可解释性——尤其是特征级可解释性，它能阐明单个特征对预测的贡献。例如，在工业场景中，有效的预警不仅需要异常告警，还需要准确识别出表明潜在故障的传感器读数，从而进行针对性维修，而不是大规模拆卸。

缺乏这种特征级可解释性会削弱用户对时间序列模型输出的信任，并使开发者的调试更加复杂，从而严重阻碍其在预警应用中的实际部署，并引发严峻的法律、伦理和公平性问题 [14, 11, 15, 16, 17, 18]。金融领域可解释性不足的一个突出例子是2010年的“闪崩”：美国股市在数分钟内暴跌近1000点，其触发原因是交易系统基于无法实时解释的时间序列模式执行了大规模抛售。该事件抹去了约1万亿的市值，最终削弱了人们对算法交易系统的信心。

鉴于可解释性对基于时间序列的早期预警系统至关重要，人们已投入大量精力开发可解释的时间序列预测方法。这些方法包括与深度学习架构相结合的方法，例如用于卷积神经网络的基于反向传播的方法 [19]和基于扰动的方法 [20]，以及用于循环神经网络的注意力机制[21], ，还有专门为时间序列预测设计的模型，如经典的ARIMAX模型 [22], ，以及更新近的基于符号聚合近似（SAX）的方法 [23], 、模糊逻辑方法 [24], 和基于形状片段的模型 [25]。此外，模型无关的可解释性技术，包括针对时间序列数据对LIME和SHAP的扩展，也为这一不断发展的领域做出了贡献 [26]。关于可解释时间序列预测方法的全面综述可供进一步参考 [10, 26]。

尽管取得了这些进展，现有时间序列可解释性方法与基于时间序列的早期预警应用的实际需求之间仍存在显著差距。

这一差距主要体现在三个方面：（1）事后解释忽视时间依赖关系：时间序列数据具有很强的时间依赖关系 [27], ，而早期预警应用需要解释能够阐明特定时间特征如何影响最终预测。然而，现有事后方法往往无法满足这一要求，因而给出的解释并不可靠：基于 LIME 的方法会针对每个数据点分别构建替代模型 [28], ，而基于 SHAP 的方法则会独立评估相邻时间步的特征 [29]。（2）深度学习方法的特征级可解释性有限：深度学习模型的可解释性方法（例如突出显示具有影响力的子序列、注意力权重或显著时间区域的方法）主要解释输入序列中的关键信息何时出现，却无法充分揭示哪些特征驱动了预测，因此缺乏早期预警应用所需的特征级可解释性 [19, 20, 21]。此外，这些方法通常会忽略特征交互对预测结果的影响 [5]。（3）固有可解释方法存在准确性–可解释性权衡：ARI-MAX 等固有可解释模型具有特征级可解释性，但其预测准确率通常不及基于深度学习的方法 [10, 7], ，因此难以满足早期预警应用对准确性和可解释性的双重要求。

为弥补这一差距，我们提出了一种用于时间序列预测的可解释多项式学习（IPL）方法。该方法引入了显式的多项式结构，能够直接解释单个特征及其交互效应对预测的影响。我们将 IPL 与广泛使用的可解释性方法进行比较，包括 ARI-MAX、LIME 和 SHAP，并证明 IPL 不仅在识别特征贡献方面优于这些基线方法，在预测准确率和计算效率方面也表现更佳——这些属性对于实时预警应用至关重要。通过扰动分析，我们表明，对输入特征进行受控修改会引起预测结果相应变化，从而反映出受扰动特征的重要性，并验证 IPL 量化特征对预测的贡献的能力，

此外，我们在两个真实世界数据集上评估了 IPL：比特币市场数据和实地采集的天线数据。针对后者，我们构建了一种由可解释性驱动的预警机制，实验结果表明，基于 IPL 的预警机制在效率和预测准确率方面均表现出色。

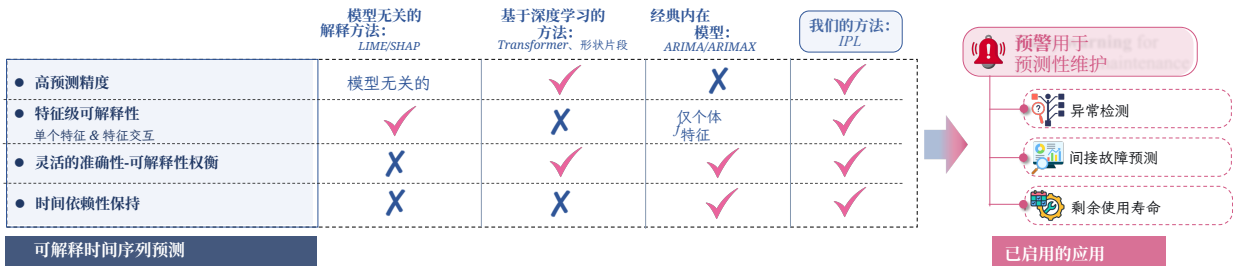


图 1：不同可解释时间序列预测方法的比较

本文主要有两项贡献。首先，我们提出 IPL 方法，为可解释时间序列预测领域的研究作出贡献。IPL 通过多项式结构对输入特征进行建模，保留时间依赖关系，并提供特征级可解释性。借助可调节的多项式阶数，IPL 能够灵活调整准确性–可解释性权衡，从而选择一个既能实现高准确率、又能提供特征级可解释性的阶数，具体落实了 [30] 提出的灵活可解释性策略。如图 1 所示，IPL 独特地融合了高准确率、适用于预警的可解释性，以及可调节的准确性–可解释性权衡。其次，我们通过证明在时间序列分析的预警领域中，除单个特征外，特征之间的交互作用对预测结果也至关重要，为该领域作出了贡献 [4] 。数值结果证实，IPL 能有效识别特征（包括单个特征和交互特征）对预测的贡献，其表现优于 ARIMAX 等固有可解释方法；这类方法仅关注单个特征。

本文其余部分安排如下。第 2 节回顾可解释时间序列预测的相关工作。第 3 节介绍所提出的可解释多项式学习方法。第 4、5 和 6 节介绍在

模拟数据、真实世界的比特币价格数据以及实地采集的天线数据上开展的实验，以验证所提出的方法。第 7 节总结全文。

2. 相关工作

在高风险应用中，建立信任、确保问责并促进模型调试的需求，推动了可解释时间序列预测日益受到重视 [31, 14, 11, 3, 32]。本综述不拘泥于事后可解释性与事前可解释性的传统二分法，而是沿着时间序列可解释性方法的发展轨迹展开，旨在厘清现有方法与预警应用实际需求之间的差距。

第一类是专为时间序列设计的经典内在可解释方法 [33, 34, 22]。ARIMAX 是一个典型示例，它通过将原始输入特征直接纳入预测模型，实现特征级可解释性。然而，这类方法仅关注单个输入特征，通常忽略特征之间的相互作用。此外，由于它们大多是线性或仅具有轻微非线性，因此往往难以捕捉现代深度学习架构能够学习的复杂模式 [10, 7], ，由此形成一种权衡：可解释性是以牺牲预测准确率为代价的。

第二类包括模型无关的事后可解释性技术，例如基于 LIME 的 [28, 35] 方法和基于 SHAP 的 [29, 36] 方法。这些技术通常应用于预训练的黑箱模型，并通过可解释代理模型近似模型行为（如 LIME），或将预测分解为特征级归因（如 SHAP）。虽然这些技术在表格数据场景中效果良好 [12] ，且能够利用复杂预训练模型的高准确率，但将其应用于时间序列时会面临显著局限。具体而言，基于 LIME 的扰动可能会破坏时间序列内在的动力学结构，从而生成不切实际的合成样本，并导致解释不可靠 [37]。同样，SHAP 将相邻时间步的特

征视为相互独立，这与时间序列固有的时间自相关性相矛盾，并削弱了解释的保真度。此外，沙普利值估计的计算成本会随着特征数量呈指数级增长，因此 SHAP 不适用于长时间序列 [29]。

第三类是基于深度学习的可解释性方法。这些方法通常基于子序列、注意力权重或显著时间区域来获得解释，[19, 20, 21], 强调时间重要性——即哪些信息至关重要，或模型在输入序列中的关注位置——而非提供原始输入特征的特征级可解释性，也不解释这些特征之间的交互如何驱动预测。在实际的预警应用中，只有基于可观察且可控的输入特征的解释，才能为决策者提供直观且可操作的洞见。

总而言之，现有的可解释时间序列预测方法存在三大局限：（1）经典线性模型的预测准确率有限；（2）模型无关方法会破坏时间结构，且生成的解释可靠性有限；（3）基于深度学习的可解释性方法主要关注衍生的时间表示，而非原始输入特征。此外，这些方法通常忽视特征交互，也缺乏在预测准确率与可解释性之间灵活权衡的能力。

为应对这些挑战，我们提出了可解释多项式学习（IPL）方法。IPL 通过多项式表示显式建模原始输入特征及其任意阶交互，将可解释性与模型结构紧密结合，从而在保留时间依赖关系的同时，为单个特征及其交互提供特征级可解释性。通过调整多项式的次数，IPL 可在预测准确率与可解释性之间灵活权衡。实验结果表明，IPL 在保持较低计算成本的同时实现了较高的准确率，因此非常适用于实时预警应用。

3. 时间序列可解释多项式学习

本节首先介绍特征级可解释性，这是我们提出该方法的动机。随后，我们阐述时间序列预测中的可解释性要求，并详细介绍可解释多项式学习方法及其针对时间序列的具体算法实现。

3.1. 特征可解释性：兼顾单个特征与特征交互

现有关于时间序列可解释性的研究主要侧重于基于单个特征的解释，却大多忽略了特征交互 [5]。然而，近期研究表明，这类解释往往无法捕捉人工智能模型为生成准确预测而利用的复杂关系。例如，在医疗保健领域，药物相互作用可能导致意料之外的协同或拮抗效应，而这些效应与患者的状况密切相关 [38]，；疾病通常也是临床因素之间复杂交互作用的结果，而非孤立的个体因素所致 [5]。因此，能够解释特征交互的方法不仅能更好地呈现预测过程，也符合领域知识。

对此，越来越多的研究致力于开发能够明确考虑特征交互的技术。一项颇具代表性的研究方向是扩展 SHAP，通过评估特征子集来纳入交互效应，从而量化特征组合的重要性 [39, 40]。一些研究尝试扩展广义加性模型（GAM），纳入成对交互项，以开发内在的、基于交互的可解释性方法 [41]，，不过，这些方法受限于 GAM 框架本身固有的加性特性。除此之外， [42] 提出了积分 Hessian 方法来评估特征交互的重要性，而 [5]则开发了一种新颖的基于图学习的模型，可自动识别患者属性之间的复杂交互及其对住院时长预测的影响。此外， [5] 对基于单个特征和基于特征交互的可解释性方法进行了比较综述，为感兴趣的读者提供了宝贵的见解。

3.2. 时间序列预测中的可解释性要求

采用 SHAP 等事后方法及相关方法，无法有效解释时间序列数据，因为这些方法会分别扰动各个时间点，从而破坏其固有的序列结构，导致解释缺乏信息，甚至产生误导。相反，我们需要一种与数据序列特性内在契合的模型，既能保留时间依赖关系，又能明确刻画跨时间特征交互（即不同时间步上的不同特征之间的交互）。因此，我们采用线性多项式模型进行时间序列预测：

$$f(x) = \sum_{m=0}^{s} \sum_{1 \le t_1 \le \cdots \le t_m \le T} \sum_{1 \le i_1 \le \cdots \le i_m \le d} w_{t_1,\ldots,t_m}^{(i_1,\ldots,i_m)} \cdot \prod_{k=1}^{m} x_{t_k}^{(i_k)}.$$

表 1 详细说明了该公式各组成部分的含义。

表 1：多项式公式各组成部分说明

组成部分	含义
s	多项式的总次数。
m	当前单项式的次数（$0 \le m \le s$）。
T	时间序列的总长度。
d	每个时间点的特征维度。
t_k	第 k 个变量的时间索引（$1 \le t_k \le T$）。
i_k	第 k 个变量的特征索引（$1 \le i_k \le d$）。
$1 \le t_1 \le \cdots \le t_m \le T$	时间索引的非递减排序。
$1 \le i_1 \le \cdots \le i_m \le d$	同一时间的特征索引按非递减排序。
$w_{t_m}^{(i_1,\ldots,i_m)\,t_1,\ldots,}$	量化该时空交互模式的重要性。
$\prod_{k=1}^{m} x_{t_k}^{(i_k)}$	由 m个变量相乘形成的单项式；$x^{(i_k)\ t_k}$ 是特征 i_k 在时间 t_k的值。

我们考虑一种特殊情况，其中 $s = 2$:

$$f(x) = w_0 + \sum_{t=1}^{T} \sum_{i=1}^{d} w_t^{(i)} x_t^{(i)} + \sum_{1 \le t_1 \le t_2 \le T} \sum_{1 \le i_1 \le i_2 \le d} w_{t_1,t_2}^{(i_1,i_2)} x_{t_1}^{(i_1)} x_{t_2}^{(i_2)},$$

其中， w_0 是常数偏置项， $w_t^{(i)}$ 用于捕捉特征 i在时间 t的重要性，而$w^{(i_1,i_2)\ t_1,\ t_2}$ 则量化特征 i_1 在时间 t_1 与特征 i_2 在时间 t_2之间的特征交互重要性。

该多项式模型的设计体现了模型结构与其可解释性之间的深度耦合。权重 $w_t^{(i)}$ 表示一种“时间–特征”效应，而 $w^{(i_1,i_2)}{}_{t_1,\ t_2}$ 则量化一种“时间–特征交互”效应，这与事后可解释性方法存在根本差异。

3.3. 可解释多项式学习模型

令 $D = \{(x_i,\ y_i)\}_{i=1}^m$ 为满足 $x_i \in \mathcal{X} \subseteq \mathbb{R}^d$ 的样本集合，并令 $y_i \in [-M, M]$ 对某个正数 M 成立。我们的目的是

$$\arg\min_{f \in \mathcal{P}_s^d} \left\{ \frac{1}{m} \sum_{i=1}^m \phi(f(x_i), y_i) \right\} \tag{1}$$

对任意 s 成立。由于 $\mathcal{P}_s^d$ 的维度为 $n = \binom{s+d}{s}$，如果我们能够选取 $\{\eta_i\}_{i=1}^n \subset \mathbb{R}^d$，使得 $\{(1 + \eta_i \cdot x)^s\}_{i=1}^n$ 是一个线性无关系统，那么

$$\mathcal{P}_s^d = \left\{ \sum_{i=1}^n c_i (1 + \eta_i \cdot x)^s : c_i \in \mathbb{R} \right\} =: \mathcal{H}_{\eta,n}, \tag{2}$$

其中 $\mathcal{H}_{\eta,\ n}$ 是为求解（1）而构造的有限基线性空间。通过这种方式，（1）可以转化为

$$\arg\min_{f \in \mathcal{H}_{\eta,n}} \left\{ \frac{1}{m} \sum_{i=1}^m \phi(f(x_i), y_i) \right\}. \tag{3}$$

为了验证（3）的可行性，我们需要做两件事。其一是给出 $\{\eta_i\}_{i=1}^n$ 的选择策略，使（2）成立；其二是确保矩阵 $A_{m,\ n} := ((1 + x_i \cdot \eta_j)^s)_i^{m\ n}{}_{,\ j=1}$ 的非奇异性。

我们首先基于 Haar空间和基本系统的概念，介绍针对 $\{\eta_i\}_{i=1}^n$ [43]的选择策略 [43], 。根据 $\{\eta_i\}_{i=1}^n$，如果可以在 $\mathbb{R}^d$ 中选取一组点 $\mathcal{H}_{\eta,\ n}$ 使其构成一个 $(n+1)$ 维的 Haar空间，那么上述所有问题都可以得到解决。然而，对于 $d \geq 2$，在 $\mathbb{B}^d$ 上不存在维度为 $N \geq 2$ [44]的 Haar空间。在此情形下，我们采用 K_s 中介绍的与多项式核 [44]相关的基本系统。

定义1。 令 $\zeta := \{\zeta_i\}_{i=1}^n \subset \mathbb{R}^d$。 ζ 称为一个 K_s-基本系统，如果 $dim\mathcal{H}_{\zeta,\ n} = \binom{s+d}{s}$.

根据上述定义，容易看出，任意的 K_s-基本系统都蕴含条件 (2)。此外， [44, 命题 1] 证明，几乎所有的 $n = \binom{d+s}{s}$ 点都构成一个K_s-基本系统。

根据 [45, 命题1], ，我们可以设计简单的策略来选择中心点 $\{\eta_j\}_{j=1}^n$。特别地，可以按照均匀分布独立同分布地随机选择 $\{\eta_j\}_{j=1}^n$ ，因为相对于勒贝格测度，均匀分布是连续的 [46]。那么，几乎必然有

$$\mathcal{P}_s^d = \left\{ \sum_{i=1}^n c_i (1 + \eta_i \cdot x)^s : c_i \in \mathbb{R} \right\}.$$

根据 [44], ，由于 n 通常远小于 m，矩阵 $A_{m,\ n}$ 几乎必然以接近 1 的概率非奇异，这证明了模型 (3) 的可行性。

3.4. 时间序列的可解释多项式学习算法

基于上一节介绍的 IPL 模型，我们现在介绍其针对时间序列预测的改编版本。在此之前，我们先说明时间序列数据的特点。与独立同分布（i.i.d.）样本不同，时间序列中的观测值天然存在时间依赖性。具体而言，时间依赖性不仅存在于输入与输出之间，也存在于不同时间步的输入之间以及不同时刻的输出之间。因此，除了纳入单个输入特征及其交互项（如公式（2）所示）之外，时间序列 IPL 算法还必须灵活地将时间信息嵌入模型中。具体来说，模型输入扩展为包含滞后变量，即输入特征和目标标签的过去观测值。为了满足预警应用的需求——在这些应用中，用户

通常关注影响最大的特征，我们引入可解释性阈值，仅筛选出最相关的项，从而得到一个稀疏且可解释的模型。

根据算法1，对于给定的数据集，IPL 只需指定多项式次数 s、滞后输入变量的数量（L_x）、目标标签的数量（L_y）以及可解释性阈值 I。基于这些设置，该算法会生成一个可解释估计器，同时量化各个特征及其交互作用对每个预测的贡献。所提出的算法具有以下优势：

- 原生支持时间序列预测：IPL 的特征级可解释性与其模型架构天然紧密相连。它不仅能够适应时间序列数据固有的时间依赖关系，还支持灵活地将滞后变量作为输入。这种设计符合时间序列预测中广泛采用的实践 [3, 16]。

- 灵活权衡准确性与可解释性：只需调整多项式次数，IPL 即可在准确性与可解释性之间实现平衡。不同于需要大量超参数调优和领域专业知识的深度学习模型，IPL 的简洁性使非专业人士也能轻松使用，从而促进其在预警中的实际应用。

- 预警应用中的低计算开销：IPL 采用交替方向乘子法进行优化。针对各种损失函数，已有高效的基于交替方向乘子法的求解策略，可确保模型的可扩展性，并便于将其应用于实时预警应用。

注1. 对于 *IPL* 算法的第 *4* 步，u^{k+1} 的更新可解析地表示为

$$u^{k+1} = \left(\beta A^\top A + \alpha I_n\right)^{-1} \left[\alpha u^k + \beta A^\top v^k - A^\top w^k\right],$$

算法1 时间序列可解释多项式学习

1. 输入：样本集 $D=\{(x_i,y_i)\}_{i=1}^{T}$，其中 T 表示最终数据采集时间戳，$x_i\in\mathbb{R}^d$，多项式核 $K_s(x,x')=(1+x\cdot x')^s$，多项式次数 $s\in\mathbb{N}$，时间滞后参数 L_x,L_y，以及可解释性阈值 I。

2. 确定核中心： 设 $n=\binom{s+d}{s}$（d 个变量上次数为 s的多项式空间的维度），并取 $\{n_j\}_{j=1}^{n}$，以$n_j=x_j$ 作为 K_s 基本系统。

3. 构造新时间序列：定义 $\tilde{x}_i=(x_i,x_{i-1},\ldots,x_{i-L_x},y_{i-1},\ldots,y_{i-L_y})\in\mathbb{R}^{d+L_x+L_y}$,$0\le L_x<i$, $0\le L_y<i$，从而得到扩展样本集 $\tilde{D}=\{(\tilde{x}_i,y_i)\}_{i=1}^{T}$。形成核矩阵 $A:=\left(K_s(\tilde{x}_i,n_j)\right)_{,i=1,j=1}^{T}$.

4. 用交替方向乘子法求解目标函数 (3)：通过引入另一个变量 v，将 (3) 改写为以下优化问题，

$$\min_{u\in\mathbb{R}^n,v\in\mathbb{R}^T} f(v)\quad \text{s.t.}\quad Au-v=0, \tag{4}$$

其中 $f(v):=\frac{1}{T}\sum_{i=1}^{T}\phi(v_i,y_i)$。式 (4) 的增广拉格朗日函数定义为

$$\mathcal{L}_\beta(u,v,w)=f(v)+\langle w,Au-v\rangle+\frac{\beta}{2}\|Au-v\|_2^2,$$

其中 $w\in\mathbb{R}^T$ 是乘子变量，$\beta>0$ 是增广拉格朗日参数。基于 $\mathcal{L}_\beta$，式 (3) 的交替方向乘子法算法如下：给定初始化 u^0,v^0,w^0，参数 $\alpha>0,\beta>0$，对于 $k=0,1,\ldots,$

$$u^{k+1}=\arg\min_{u\in\mathbb{R}^n}\left\{\mathcal{L}_\beta(u,v^k,w^k)+\frac{\alpha}{2}\|u-u^k\|_2^2\right\}, \tag{5}$$

$$v^{k+1}=\arg\min_{v\in\mathbb{R}^T}\mathcal{L}_\beta(u^{k+1},v,w^k), \tag{6}$$

$$w^{k+1}=w^k+\beta(Au^{k+1}-v^{k+1}). \tag{7}$$

采用近端更新策略计算式 (5) 中的 u^{k+1}，根据具体的损失函数 ϕ 更新 v^{k+1}，最终得到式 (3) 的全局极小化解：$f_{\tilde{D},s}(\tilde{x})=\sum_{j=1}^{n}u_j^*K_s(n_j,\cdot)$。

5. 提取特征重要性：将预测器 $f_{\tilde{D},s}(\cdot)$ 表示为定义在 $(d+L_x+L_y)$ 维输入 $\tilde{x}$ 上的显式多项式展开：

$$f_{\tilde{D},s}(\tilde{x})=\sum_{|\boldsymbol{\alpha}|\le s}\omega_{\boldsymbol{\alpha}}\,\tilde{x}_1^{\alpha_1}\cdots\tilde{x}_{d+L_x+L_y}^{\alpha_{d+L_x+L_y}},$$

其中，$\boldsymbol{\alpha}=(\alpha_1,\ldots,\alpha_{d+L_x+L_y})$ 是一个多重指标，$\alpha_j\in\mathbb{N}_0$ 表示第 j 个变量的指数，而 $|\boldsymbol{\alpha}|:=\sum_{j=1}^{d+L_x+L_y}\alpha_j$ 表示多重指标的阶（当 $|\cdot|$ 应用于标量时，表示绝对值）。系数 $\omega_{\boldsymbol{\alpha}}$ 量化了相应单项式项对预测的贡献。

6. 选择特征和系数：将保留下来的多重指标 $\boldsymbol{\alpha}$ 及其对应的系数 $\omega_{\boldsymbol{\alpha}}$ 收集到一个稀疏集 $\Omega_I:=\{(\boldsymbol{\alpha},\ \omega_{\boldsymbol{\alpha}})\mid|\boldsymbol{\alpha}|\le s,\ |\omega_{\boldsymbol{\alpha}}|\ge I\}$.

7. 输出：可解释预测器 $f_{\tilde{D},s}^{I}(\tilde{x})=\sum_{(\boldsymbol{\alpha},\omega_{\boldsymbol{\alpha}})\in\Omega_I}\omega_{\boldsymbol{\alpha}}\,\tilde{x}_1^{\alpha_1}\cdots\tilde{x}_{d+dL_x+L_y}^{\alpha_{d+dL_x+L_y}}$，及其稀疏系数集 Ω_I。

其中， I_n 是大小为 n 的单位矩阵。对于平方损失，(3)的解析解为 *[44]*：$f_{\tilde{D},s}(\cdot) = \sum_{j=1}^{n} c_j K_s(n_j, \cdot)$，其中 $\mathbf{c} = (c_1, \ldots, c_n)^\top$ 由 $\mathbf{c} = \mathrm{pinv}(A)\mathbf{y}$给出。此处，$\mathrm{pinv}(\cdot)$ 表示伪逆算子；而对于合页损失， *[45]*中提供了一种快速迭代求解器。

若 ϕ 可微，则 v更新可分解为 T 个相互独立的标量方程：

$$\frac{1}{T}\phi'(v_i, y_i) + \beta(v_i - b_i) = 0,$$

其中 $b_i = (Au^{k+1})_i + w_i^k/\beta$。对于凸损失函数（如平方损失、逻辑损失或合页损失），每个方程都对应于最小化一个一维凸函数。因此，可以通过牛顿法高效求解：对于每个 i,

$$v_i^{(t+1)} = v_i^{(t)} - \frac{\frac{1}{T}\phi'(v_i^{(t)}, y_i) + \beta(v_i^{(t)} - b_i)}{\frac{1}{T}\phi''(v_i^{(t)}, y_i) + \beta},$$

通常只需经过几次迭代即可收敛。

具体而言，对于逻辑损失 $\phi(v, y) = \log(1 + e^{-yv})$，其中 $y \in \{-1, 1\}$，牛顿迭代为

$$v_i^{(t+1)} = v_i^{(t)} - \frac{\frac{y_i}{T}\left[\sigma(y_i v_i^{(t)}) - 1\right] + \beta(v_i^{(t)} - b_i)}{\frac{1}{T}\sigma(y_i v_i^{(t)})\left[1 - \sigma(y_i v_i^{(t)})\right] + \beta},$$

其中 $\sigma(z) = 1/(1 + e^{-z})$。

4. 平方损失下的模拟数据分析

本节中，我们开展了三组实验，以评估所提出的 IPL 在平方损失函数下的有效性。首先，我们探究时变目标变量在时间序列分析中的重要性。其次，我们展示 IPL 在特征可解释性、预测准确率和计算效率方面的优势。最后，我们通过扰动各可解释性方法识别出的最重要特征进行扰动分析，以评估哪种方法最准确地识别出对准确预测至关重要的特征。

该可解释　本研究中使用的可解释性方法分为两类：（1）

事后可解释性方法：SHAP 和 LIME。对于 SHAP，使用深度神经网络（DNN）计算特征贡献；而在 LIME 中，核岭回归（KRR）用作局部代理模型。（2）事前可解释性方法：ARIMAX。SHAP、LIME 和 ARIMAX 的详细实验配置见附录。除非另有说明，所有后续实验均采用这四种方法的相同设置。具体而言，根据数值验证中的最优表现，IPL 的多项式次数 s 设为 2。

我们生成由训练样本 $\{(x_t, y_t)\}_{t=1}^{T}$ 和测试样本 $\{(x_{t'}^*, y_{t'}^*)\}_{t'=1}^{T'}$组成的时间序列数据。协变量 x_t 和 $x_{t'}^*$ 独立同分布地取自 5 维超立方体 $[0,1]^5$，并服从均匀分布。该序列的时间演化由目标函数 $y_{t+1} = \cos(y_t)\sin(y_{t-1}) + f(x^t) + \epsilon^t, y_{t'+1}^* = \cos(y_{t'}^*)\sin(y_{t'-1}^*) + f(x_{t'}^*)$, 描述，函数 $f(\cdot)$ 定义如下，其中 $\epsilon^t \sim \mathcal{N}(0, 0.1_2)$

$$f(\cdot) = (1 + x_1^t + 2x_2^t + 3x_3^t + 4x_4^t + 5x_5^t + 6x_1^t x_2^t - 7x_3^t x_4^t)/15.$$

训练样本量设为 $T = 4,000$，测试样本量设为 $T' = 1,000$。所有实验结果均取自 10 次独立试验的平均值。除使用图形处理器进行的 LIME 相关实验外，其他实验均使用 Python 3.7，在配备 Intel Core i5 2GHz 处理器的 PC 上完成。用于复现实验的脚本可在 `https://github.com/Ariesoomoon/IPL_TS_experiments`获取。

4.1. 时间序列中时间依赖关系的重要性

图2展示了整合滞后目标变量对回归性能的影响。这里，“-x6”表示将原始特征与前一个时间滞后的目标变量相结合后形成的6维输入，“-x7”表示将原始输入与前两个时间滞后的目标变量相结合，依此类推。结果表明，随着纳入的滞后变量增多，预测准确率持续提升，凸显了在时间序列

预测中纳入历史目标信息的价值。然而，从 x_{13}开始，我们观察到预测准确率趋于稳定。这可能是因为引入冗余特征所带来的负面影响，开始抵消纳入更多滞后目标变量所带来的预测价值。

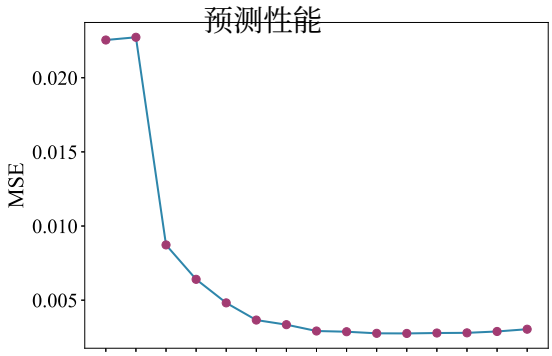


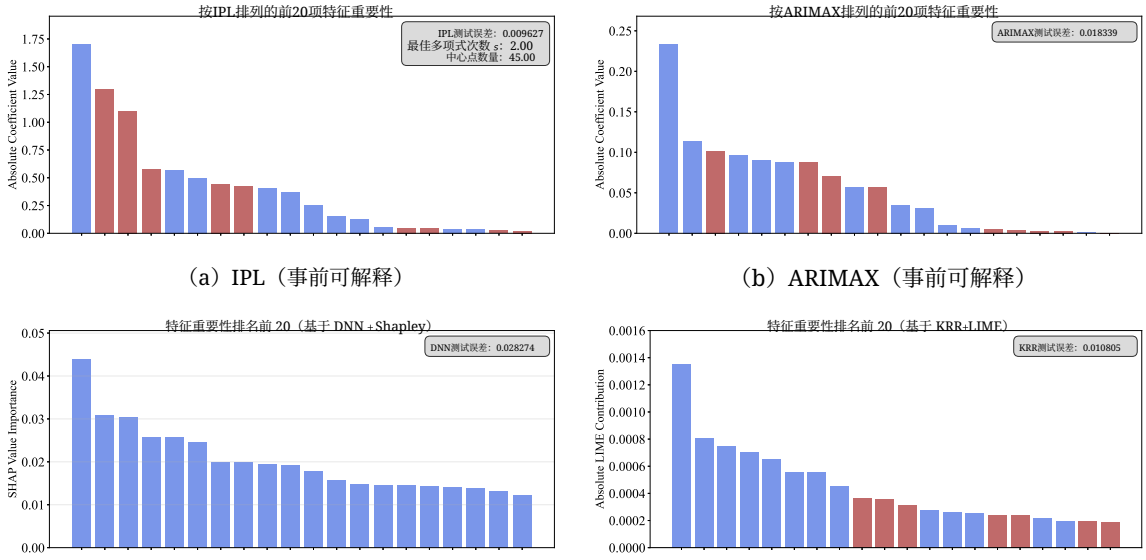
图2：时间相关目标变量的重要性（“-x6”表示将原始特征与一个滞后目标变量组合形成的6维输入，“-x7”表示与两个滞后目标变量组合形成的输入，以此类推）。

4.2. IPL 在提供特征可解释性方面的优势

本实验将 IPL 与 SHAP、LIME 和 ARIMAX 进行比较，以展示 IPL 的优势。我们引入两个滞后目标变量 x_7，x_8，以及一个无关输入变量 x_6，以便更可信地考察各方法排除该变量的能力。

图3展示了不同可解释性方法识别出的排名前20的重要特征。为确保对其特征可解释性能力进行稳健比较，特意将 IPL 的预测准确率维持在最低水平。可以得出三个关键观察结果：(1) 只有 IPL 得出的特征重要性排名和大小与目标函数完全一致。ARIMAX 在这方面也表现出色，但在实践中其训练过程较慢，因为它需要通过网格搜索确定最优的 $(p,\ d,\ q)$ 阶数。(2) IPL 能有效降低无关特征x_6 的优先级（将其排在第 17^{th} 位）。虽然 LIME 也能类似地识别无关变量，但它需要耗费更多的计算时间和内存。事实上，

本实验中的 LIME 结果是在 GPU 加速下获得的，其运行时间约为在 CPU 上运行的 IPL 的 1000 倍。(3) 所有方法，尤其是 IPL 和 ARIMAX 这类内在可解释的方法，都成功识别出了所纳入的滞后目标变量的重要性。



（a）IPL（事前可解释）　（b）ARIMAX（事前可解释）

(c) DNN+SHAP（事后）　(d) KRR+LIME（事后）

图3：不同可解释性方法识别出的最重要特征（注：横轴上属于目标函数的特征以红色突出显示；图中的红色和蓝色柱分别代表负值和正值）。

表2：不同评估指标下可解释性方法的性能比较

指标	增加三个输入变量后的原始输入 x_6，x_7，x_8(维度 $=8$)				仅原始输入 (维度 $=5$)			
	IPL	SHAP	LIME	ARIMAX	IPL	SHAP	LIME	ARIMAX
特征可解释性								
特征重叠比	0.7143	0.4286	0.1429	0.7143	1.0000	0.4286	0.7143	0.8571
排序相似度	1.0000	0.6699	0.7881	0.8929	1.0000	0.8571	0.6847	0.4324
数值相似度	0.9979	0.1734	0.4723	0.9835	0.9992	0.2759	0.2831	0.8487
前10个特征上的预测准确率（均方误差）								
线性回归（LR）	0.0064	0.0073	0.0096	0.0064	0.0159	0.0168	0.0165	0.0158
效率（运行时间/每次试验）	1.81 秒	32.45 秒	32.50 分钟	20.57 秒	1.47 秒	20.55 秒	17.03 分钟	5.56 秒

注：每个指标的最佳结果以彩色突出显示。 *Feature overlap ratio* 表示目标函数中的特征同时包含在该方法前十特征集中的比例。 *Ranking similarity* 定义为 $\rho = 1 - (6\sum_{i=1}^{n} d_i^2)/(n(n^2-1))$，其中 $d_i = R_{\text{method}}(i) - R_{\text{target}}(i)$表示该方法的排序 R_{method} 与真实排序 R_{target}中第 i个特征的排名差异。 *Value similarity* 定义为$\frac{\mathbf{a}\cdot\mathbf{b}}{\|\mathbf{a}\|\|\mathbf{b}\|} = (\sum_{i=1}^{n} a_i b_i)/(\sqrt{\sum_{i=1}^{n} a_i^2}\sqrt{\sum_{i=1}^{n} b_i^2})$，其中$\mathbf{v} = (v_1, \ldots, v_n)$表示跨越 n 个目标特征的向量。“*sec*”表示秒，“*min*”表示分钟。

接下来，我们将从三个关键维度对不同方法进行全面比较：f

特征可解释性、预测准确率和计算效率。

如表2所示，IPL在所有评估方面均持续优于SHAP和LIME。具体而言，它通过识别与目标函数中定义的真实特征重要性高度一致的特征，提供了更忠实的解释；使用其识别出的前十个重要特征时，可获得最佳预测准确率；并且在测试阶段所需的运行时间最短。为直观概括IPL的优势，我们将表2中的结果绘制成雷达图。如图4所示，IPL以红色阴影区域表示，占据的面积最大，清楚地展现了它在同时实现高可解释性、准确率和效率方面的整体优势。

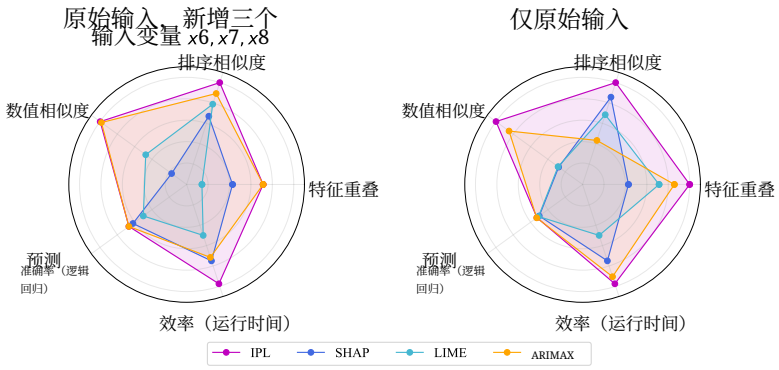


图4：比较可解释性方法性能的雷达图（越接近外边界，表示该维度上的能力越强）。

4.3. 扰动分析

我们采用了基于与广泛使用的置换特征重要性方法相同原理的扰动分析 [47]，以验证不同可解释性方法所识别的特征重要性。扰动后预测性能显著下降，表明该特征确实具有重要作用，从而验证了相应可解释性方法的有效性。具体而言，对于每种方法中最重要的特征，我们添加高斯噪声 $\epsilon \sim \mathcal{N}(0,\ \sigma^2)$对其进行扰动，其中 σ 表示该特征的标准差，同时保持其他所有特征不变。随后，我们比较进行10次试验前后的测试均方误差，并始终使用线性回归作为估计器，以确保公平比较。

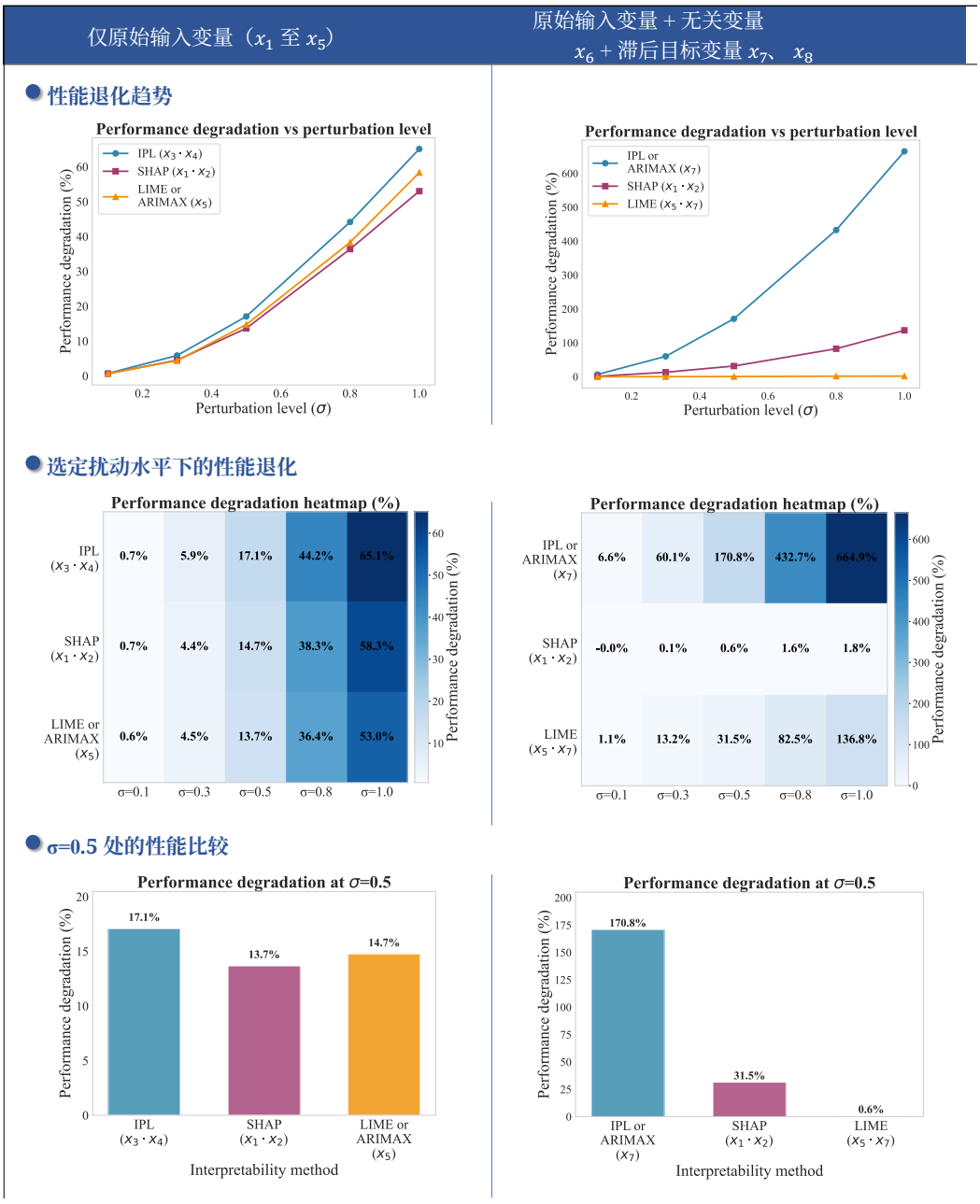


图5：扰动分析（每个方法选出的最重要特征标示在括号中）。

结果如图5所示，左侧展示原始数据上的扰动分析，右侧展示添加一个无关输入变量（x_6）和两个滞后目标变量（x_7 和 x_8）作为输入变量时的扰动分析。为直观突出性能变化的趋势和幅度，结果结合曲线和热力图进行呈现。我们看到，对最重要的

IPL 识别出的重要特征在所有扰动水平下都会导致最显著的性能退化，展现了 IPL 在特征可解释性方面的优势。纳入滞后目标变量后，这一优势更加明显。我们进一步通过柱状图比较了三种方法在扰动水平为$\alpha = 0.5$时的表现。结果表明，加入滞后目标变量后，IPL 的最大性能退化达到 170.8%，显著高于 SHAP 的 31.5% 和 LIME 的 0.6%。

5. 合页损失下的比特币历史价格数据分析

在金融时间序列分析中，预测不同预测时域下的价格走势模式十分重要，因为这会直接影响交易决策和盈利能力。本研究使用了从2024年1月1日至今的比特币历史价格数据[4]。该数据集包含开盘价、最高价、最低价、成交量和收盘价等属性。我们将开盘价、最高价、最低价和成交量作为输入，收盘价作为输出。为确保计算效率并保持时间连续性，我们从原始高频数据集中按固定间隔系统地抽取了约5,000个数据点，同时保留观测值的时间顺序。

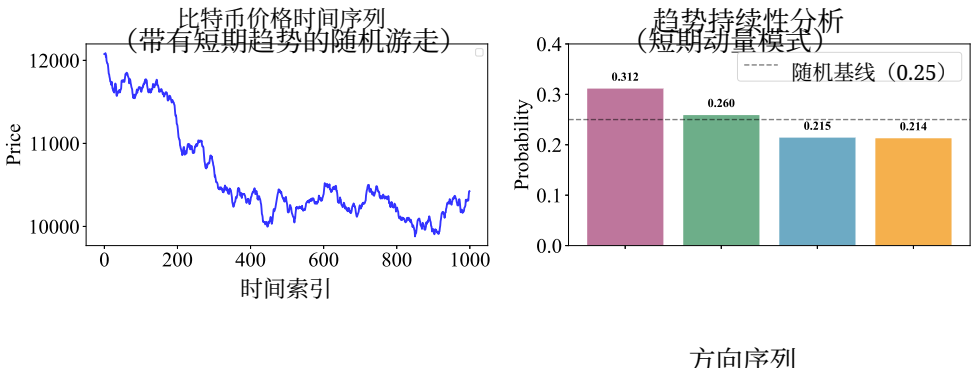


图6：比特币价格动态分析

[4]https://www.kaggle.com/datasets/mczielinski/bitcoin-historical-data/data

我们首先考察价格变化的方向模式。如图6所示，“Up $\rightarrow$ Up”表示某一时点出现上升趋势后，下一时点再次出现上升趋势。分析结果表明，“Up $\rightarrow$Up”和“Down $\rightarrow$ Down”这两种情况出现的概率均显著高于随机基线（0.25），这说明存在短期动量——也就是说，价格在此前上涨后往往继续上涨，在此前下跌后往往继续下跌。基于这一观察，我们构建了两个二元分类任务，用于对不同预测时域下的价格走势建模，并在模型训练中采用合页损失函数。目标变量标记为 `Target` k`period`，定义于 $k \in \{3, 5\}$，其中 k 表示预测时域。每个目标变量均采用二元方案对价格方向进行编码：1表示价格上涨，而 -1 表示经过 k 个时间单位后价格下跌。目标变量的数学定义如下：

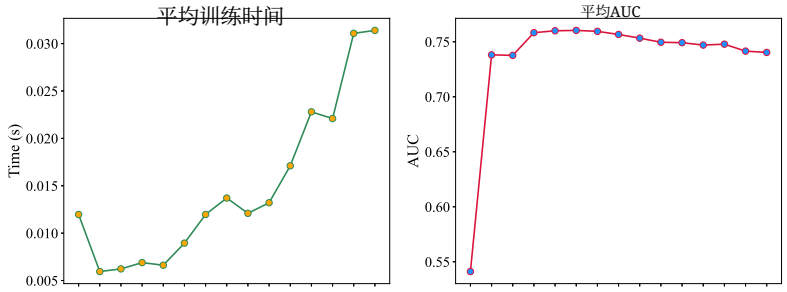
-

$$\text{Target_}k\text{period} = \begin{cases} 1 & \text{if } P_{t+k} > P_t \\ -1 & \text{otherwise} \end{cases} \tag{5}$$

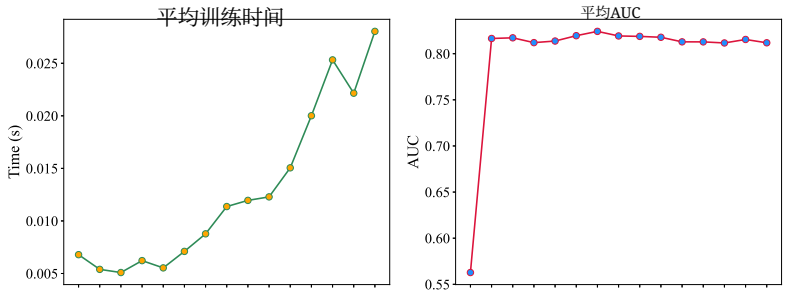
其中， P_t 表示时间 t的收盘价。

5.1. 时间序列中时间依赖关系的重要性

图7显示，在两种预测时域下，平均AUC均先快速提升，随后略有下降，这证实了历史价格信息在价格方向分类中的价值。值得注意的是，这种准确率的提升并未增加计算成本，体现了IPL的计算效率。



(a) 目标3周期



(b) 目标5周期

图7：滞后时间相关目标变量的重要性

表3：目标3周期的分类性能比较

特征	方法	类别*1*性能			准确率
		精确率	召回率	*F1* 分数	
前三名	LIME	0.73	0.51	0.60	0.64
	SHAP	0.76	0.56	0.65	0.69
	ARIMAX (2,0,2)	0.58	0.53	0.55	0.53
	IPL	0.75	0.74	0.74	0.73
第一名	LIME	0.72	0.71	0.71	0.70
	SHAP	0.74	0.55	0.63	0.68
	ARIMAX (2,0,2)	0.55	0.51	0.53	0.50
	IPL	0.75	0.74	0.74	0.73

注：$Precision = \frac{TP}{TP+FP}$、$Recall = \frac{TP}{TP+FN}$、$F1\text{-}score = \frac{2 \times Precision \times Recall}{Precision+Recall}$ 和 $Accuracy = \frac{TP+TN}{TP+FN+TN+FP}$，其中 TP、TN、FP和 FN分别表示真正例、真负例、假正例和假负例。最佳结果以颜色突出显示。

*5.2. IPL*在识别重要特征方面的优势

表3比较了基于不同可解释性方法所识别的特征构建的常用可解释决策树模型，在加密货币价格数据 [1] 上使用通用指标进行评估后的表现。结果表明，使用IPL识别出的

无论采用前三个特征，还是仅采用最重要的单个特征，所选特征的准确率始终高于使用 LIME、SHAP 或 ARIMAX 识别出的特征。这些结果表明，IPL 能够提供更可靠的特征级可解释性。

6. 现场实验：天线健康管理

天线维护人员持续监测天线系统和环境传感器的数据，以评估运行状态并进行异常检测，从而实现预测性维护。因此，亟需开发一种基于机器学习的早期预警机制，在实现高预测精度的同时保持特征级可解释性。本研究使用的数据集包含从中国电子科技集团公司第三十九研究所的实体天线设备采集的 5,437 个样本。数据按时间顺序划分为训练集（3,437 个样本）、验证集（1,000 个样本）和测试集（1,000 个样本），其中最新的样本划入测试集。表 4 对数据集中的变量进行了详细说明。

表4：现场采集的天线数据说明

变量	描述	范围/备注
x_1	位置角度值	$0^\circ \sim 360^\circ$，通过组合 x_4 和 x_5得到
x_2	速度	$0 \sim 2000$ 转/分钟，由转速计采集
x_3	电流比	$0 \sim 100\%$，采集电流与额定电流之比
x_4	粗码值	$0 \sim 65535$，16 位二进制，旋变反馈
x_5	细码值	$0 \sim 65535$，16 位二进制，旋变反馈
x_6	标签	-1 =正常，1 =异常

注：鉴于角变量 x_1 的周期性，将其转换为三角函数表示：$x_{1\sin}$ =sin($2\pi x_1$) 和 $x_{1\cos}$ =cos($2\pi x_1$)。

图8展示了数据集的初步可视化结果，突出了两个关键发现。首先，异常标签呈现出明显的周期性模式，并且往往会在特定时间区间内连续出现。其次，只有位置角和粗码值能够清晰地区分正常与异常状态。然而，无论是

单独使用位置角还是粗编码，都不足以实现准确的预警，而其他特征缺乏区分能力，这凸显了分析更高阶特征交互以提升预警效果的必要性。

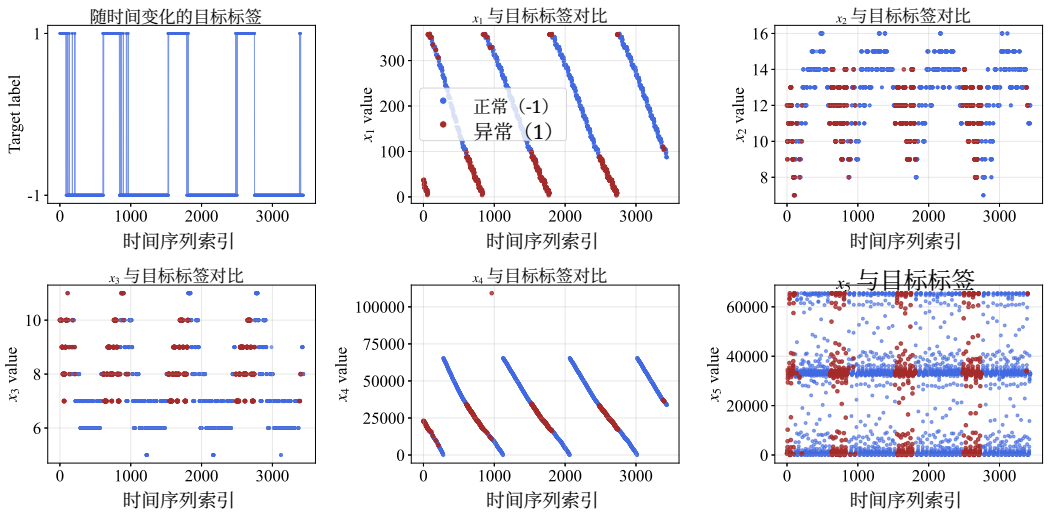


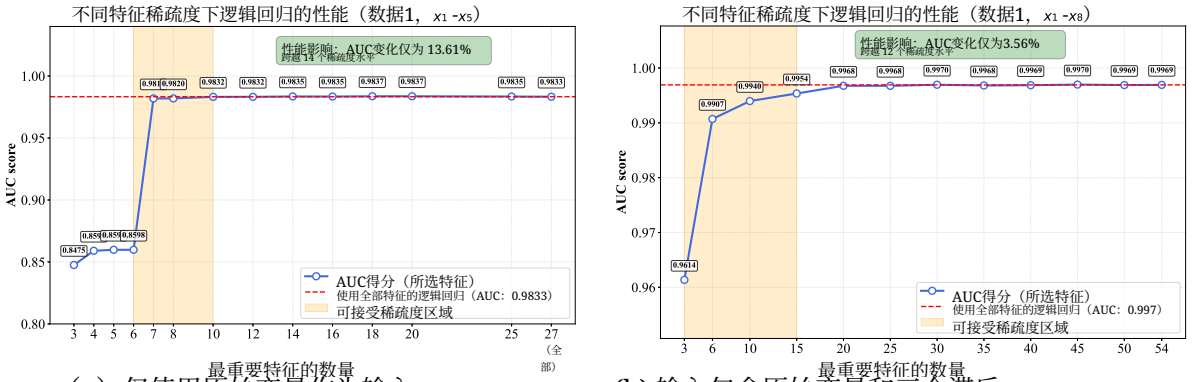
图8：不同特征值下目标标签的分布

我们在合页损失下进行了两组实验。首先，我们通过考察特征稀疏度与预测准确率之间的关系，验证IPL在提供特征可解释性方面的有效性。其次，我们比较不同可解释性方法在设计早期预警机制中的表现，以展示IPL的优势。为公平比较，所有可解释性方法均先被调优至达到相近的高准确率（> 0.95）。鉴于实验设置与可视化方式与先前实验相似，相应的特征重要性排名收录于附录中供参考。

6.1. 用于评估可解释性的特征稀疏度-准确率分析

本实验探究 IPL 选出的排名靠前的特征数量与最终分类准确率之间的关系，以评估 IPL 在提供特征可解释性方面的有效性。为确保比较公平，所有稀疏度水平下的预测均使用逻辑回归分类器。结果如图9所示，其中包括两种输入条件下的表现：仅使用最初采集的特征，以及将时间相关目标变量作为输入的一部分。

浅黄色区域表示约简特征集仍能保持较高准确率的稀疏度范围。根据该区域，我们得出结论：数据采集者只需关注排名靠前的 7 到 10 个特征，即可获得与完整特征集相当的分类准确率。此外，不同稀疏度水平下 AUC 的变化幅度极小，最大波动仅为 13.61%。这些结果证实，IPL 能有效识别并排序影响最大的特征，从而提供特征可解释性。



（a）仅使用原始变量作为输入　　(b) 输入包含原始变量和三个滞后目标标签

图9：基于 IPL 的特征稀疏度与预测性能之间的关系

6.2. 基于特征可解释性的早期预警机制设计

本节中，我们根据不同可解释性方法得出的特征重要性设计预警机制。鉴于我们的问题属于二分类问题，我们采用二叉树结构构建每种预警机制。

图10按重要性降序排列了各方法的特征，并展示了为实现令人满意的性能（每项指标均超过0.96）而设计的相应预警机制。在图中，浅灰色矩形表示对预警没有有效贡献的特征，即这些特征在正常与异常状态之间缺乏清晰的判别边界；深灰色矩形则表示构成预警规则的决策节点。每个深灰色节点下方显示了具体的预警规则。可以得出两个主要观察结果。（1）基于IPL的预警机制明显更加简单高效，仅需两个特征，且这两个特征正好是重要性排名前两位的特征。这体现了IPL在特征可解释性方面的优势：它能够准确识别对预测贡献最大的特征。（2）IPL识别出的最重要特征{v1}对应于天线转速与电流比的乘积。该交互项能够有效捕捉异常模式，因为它近似表示机械功率，而异常状态通常表现为功率异常。此外，在LIME和SHAP识别出的重要交互项中，{v2}始终存在，反映了过去状态对当前状态的影响。

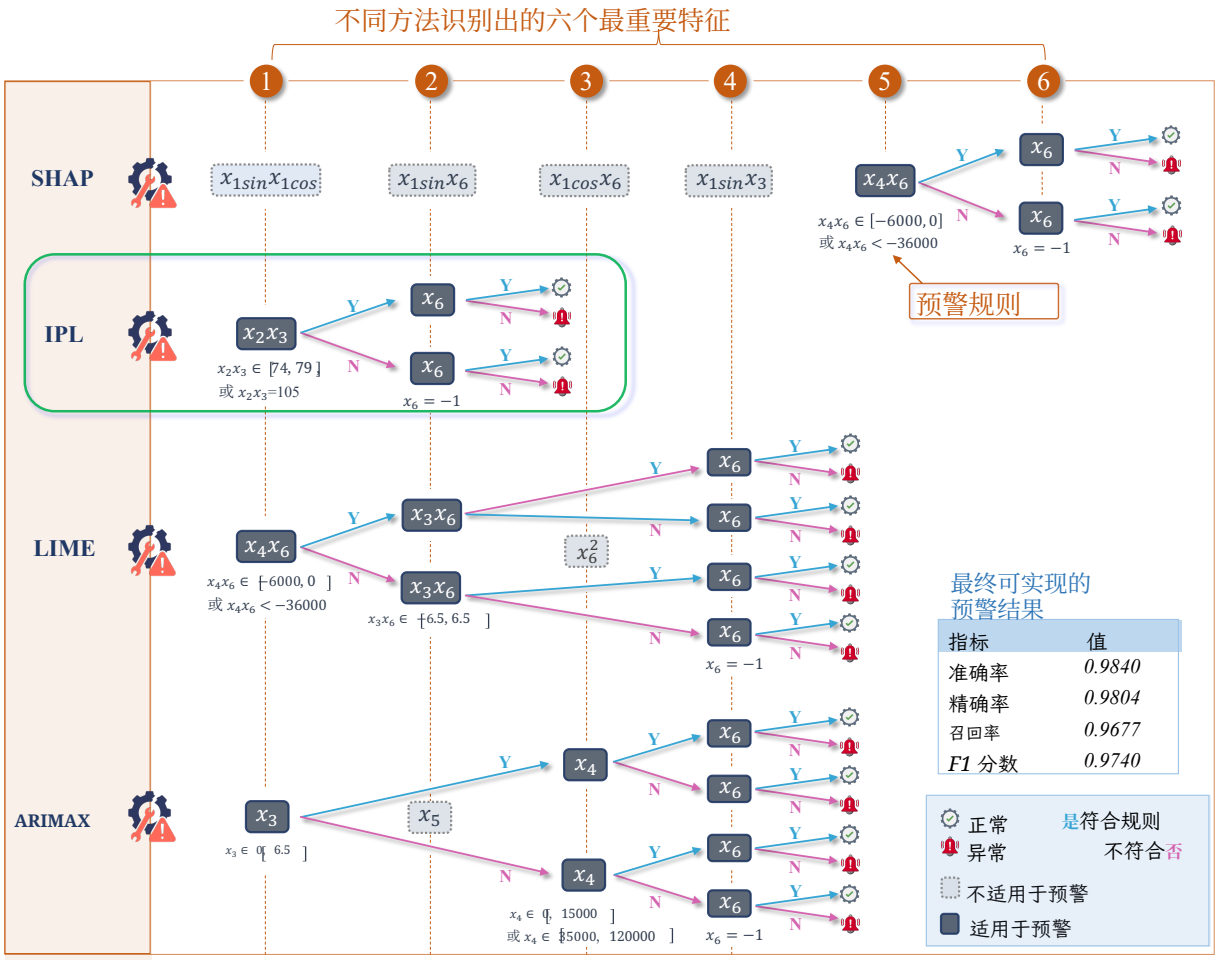
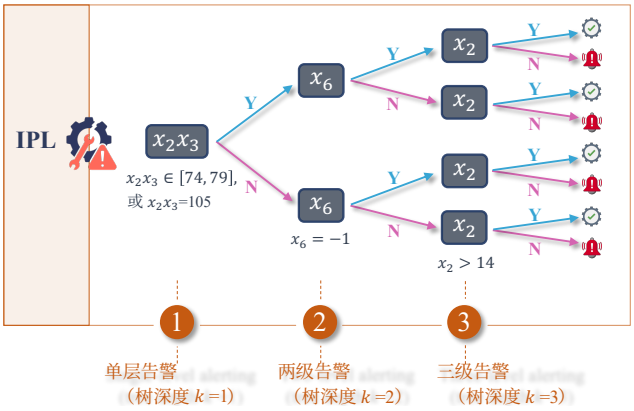


图10：基于不同方法得出的特征重要性排名设计的预警机制

警机制明显更加简单高效，仅需两个特征，且这两个特征正好是重要性排名前两位的特征。这体现了IPL在特征可解释性方面的优势：它能够准确识别对预测贡献最大的特征。（2）IPL识别出的最重要特征 x_2x_3对应于天线转速与电流比的乘积。该交互项能够有效捕捉异常模式，因为它近似表示机械功率，而异常状态通常表现为功率异常。此外，在LIME和SHAP识别出的重要交互项中， x_6 始终存在，反映了过去状态对当前状态的影响。

在天线健康管理中，除了准确率或精



	单层告警	两级告警	三级告警
指标	**值**	**值**	**值**
准确率	0.3370	0.9840	0.4390
精确率	0.3186	0.9804	0.3559
召回率	1.0000	0.9677	1.0000
F1 分数	0.4832	0.9740	0.5250

设置 k=3 可提高召回率，设置 k=2 可提高精确率

图11：基于不同树深度的 IPL 预警结果（k）

确率等传统指标外，人们更加重视对异常状态的早期检测，尤其是能够在长时间正常运行期间触发告警，以提示潜在未来异常的能力。因此，我们在图11中详细展示了基于IPL的预警机制。管理员可以采用树深度为 $k = 2$的机制来获得更高的精确率，也可以选择更深的树（$k = 3$）以实现完美召回率（1.0）。尽管后一种选择会略微降低整体精确率，但在天线维护中，这种权衡通常可以接受，因为漏检异常会导致高昂的运营成本。

尽管 $k = 2$ 机制达到了0.9677的召回率，但这一表现部分归因于天线运行状态所固有的时间连续性。通过进一步检查测试样本，我们观察到，所有对应于从正常状态到异常状态转变的实例均未被 $k = 2$ 机制检测到，但均被 $k = 3$ 机制成功捕获。表5列出了几个具有代表性的样本，进一步展示了该机制在连续未来时间步发出警告的能力，凸显了其在长期状态预测中的实际价值。

表5：设置 $k = 3$的依据：在持续正常运行期间对异常触发警报

样本索引	**预警特征**			**当前时间的标签 t_1**			**下一时间的标签 t_2**			T_n
	x_2x_3	x_6	x_2	预测值 ($k = 2$)	预测值 ($k = 3$)	真实值 真实值	预测值 ($k = 2$)	预测值 ($k = 3$)	真实值 真实值	
5034, 5048	91.0	-1	13.0	-1	1	1	-1	1	1	2
5065, 5082	96.0	-1	12.0	-1	1	1	-1	1	1	10

注： T_n 表示能够准确预测的连续未来时间步数量。 k 表示预警机制的树深度。

7. 结论

时间序列数据在实际应用中发挥着关键作用，使人们能够基于历史观测结果预测未来系统状态。最重要的应用领域之一是预警，例如工程系统中的预测性维护和医疗场景中的健康监测；在这些领域，及时的警告能够显著减少资源消耗和运营成本。因此，有必要开发高精度的预警机制，使人工操作员能够理解，并支持在原始观测变量层面进行直接干预。这促使人们需要能够同时实现高预测精度和特征级可解释性的时间序列预测方法。

现有的内在可解释方法（例如 ARIMAX）能够提供特征级解释，但通常受到预测准确率有限的制约。广泛使用的事后解释技术，如 LIME 和 SHAP，同样不适用于解释时间序列数据，因为它们往往忽略观测值之间固有的时间依赖关系。尽管已有大量基于深度学习的方法被提出，用于实现可解释的时间序列预测，但这些方法主要关注时间重要性，而不是提供原始输入变量的特征级可解释性，因此难以将其解释转化为预警所需的可操作指导。此外，这类方法通常计算开销较大，非专业人士也难以使用。总之，目前仍缺乏专门针对时间序列的预测方法，能够同时实现高预测精度和特征级可解释性，以支持预警应用。为弥补这一空白，我们提出了 IPL方法，通过多项式结构将特征与模型紧密结合，明确地将来自不同时间步的特征及其交互项纳入模型。IPL 保留了时间序列数据固有的时间结构，并可通过调整多项式阶数，在准确率与可解释性之间实现灵活平衡，使实践者能够选择一种既保持高预测精度又确保

特征级可解释性的模型。

我们通过在模拟数据和两个真实世界数据集上将 IPL 与 LIME、SHAP 和 ARI-MAX 等广泛使用的可解释性方法进行比较，验证了 IPL 在准确性–可解释性权衡方面的优势。在现场采集的天线数据集上，我们进一步比较了基于不同方法确定的特征重要性所构建的早期预警机制。结果表明，IPL 提供的可解释性能够构建更精简的早期预警机制，同时实现与其他方法相当的性能。此外，基于得出的特征重要性，可以灵活调整预警机制的深度，以满足不同的精度要求。

我们的方法适用于广泛的应用领域。例如，在金融领域，IPL 有助于识别推动大宗商品价格波动的关键因素，从而实现主动风险管理。在医疗保健领域，它可以为医务人员提供精确且可解释的健康警报，支持针对性干预，同时减轻警报疲劳。在工程设备管理领域，IPL 可协助管理人员预判维护需求并提前安排检查，从而应对日益增长的预测性维护需求。

参考文献

[1] Q. Zhang, C. Xie, Z. Weng, D. Sornette, K. Wu，广义可见曲率：一种用于加密货币泡沫识别与价格趋势预测的指标，决策支持系统185（2024年）114309。

[2] S. B. Taieb, J. W. Taylor, R. J. Hyndman，基于智能电表数据的电力需求分层概率预测，美国统计学会会刊116（2021年）27–43。

[3] N. Bertani、S. T. Jensen、V. A. Satop¨a¨a，《用于分层时间序列概率预测的联合自下而上方法》，Operations Research（2025年）。

[4] U. Mori、A. Mendiburu、I. M. Miranda、J. A. Lozano，《使用多目标优化技术对时间序列进行早期分类》，Information Sciences 492（2019年）204–218。

[5] T. Guo, I. R. Bardhan, Y. Ding, S. Zhang，采用图学习预测重症监护室住院时长的可解释人工智能方法，Information Systems Research 36 (2025年) 1478–1501。

[6] S. Hochreiter, J. Schmidhuber，长短期记忆，Neural Computation 9 (1997) 1735–1780。

[7] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez，L. Kaiser, I. Polosukhin，注意力就是你所需要的一切，神经信息处理系统进展 30 (2017年)。

[8] S. Bai，通用卷积网络和循环网络用于序列建模的实证评估，arXiv 预印本 arXiv:1803.01271 (2018年)。

[9] D. Salinas、V. Flunkert、J. Gasthaus、T. Januschowski，DeepAR：基于自回归循环网络的概率预测，《International Journal of Forecasting》36（2020年）1181–1191。

[10] B. Lim、S. Zohren，深度学习时间序列预测：综述，Philosophical Transactions of the Royal Society A 379（2021年）20200209。

[11] P. Kumar、D. Javeed、A. N. Islam、X. R. Luo，DeepSecure：一种面向网络安全决策中可解释威胁狩猎的计算设计科学方法，《决策支持系统》188（2025年）114351。

[12] J. Senoner、T. Netland、S. Feuerriegel，利用可解释人工智能改善过程质量：来自半导体制造的证据，Management Science 68（2022年）5704–5723。

[13] S. Bouazizi、H. Ltifi，利用面向卒中预测的可解释集成学习框架应用，提升基于脑电图的医疗决策中的准确率与可解释性，决策支持系统 178（2024年）114126。

.[14] B. Lim, S. O. Arik, N. Loeff, T. Pfister, 《用于可解释多时域时间序列预测的时序融合变换器》, International Journal of Forecasting 37（2021年）1748–1764。

[15] S. Wang, J. Cao, S. Y. Philip, 《用于时空数据挖掘的深度学习：综述》, 《IEEE知识与数据工程汇刊》34（2020年）3681–3700。

[16] B. Hen, N. Rabin, 《使用迭代核回归进行时空时间序列预测》, 《INFORMS数据科学期刊》4（2025年）20–32。

[17] K. Bauer, M. von Zahn, O. Hinz, 《Expl (AI) ned：可解释人工智能对用户信息处理的影响》, Information Systems Research 34（2023年）1582–1602。

[18] K. Bauer, A. Gill, 《墙上的镜子，告诉我：算法评估、透明度与自我实现预言》, Information Systems Research 35 (2024年) 226–248。

[19] F. Oviedo, Z. Ren, S. Sun, C. Settens, Z. Liu, N. T. P. Hartono, S. Ramasamy, B. L. DeCost, S. I. Tian, G. Romano 等, 《利用数据增强和深度神经网络对小型X射线衍射数据集进行快速且可解释的分类》, 《npj计算材料》5 (2019年) 60。

[20] K. Kashiparekh, J. Narwariya, P. Malhotra, L. Vig, G. Shroff, 《ConvTimeNet：用于时间序列分类的预训练深度卷积神经网络》, 载于：2019年国际神经网络联合会议（IJCNN）, IEEE, 2019年, 第1–8页。

[21] E. Choi、M. T. Bahadori、J. Sun、J. Kulas、A. Schuetz、W. Stewart, 《Retain：一种利用反向时间注意力机制的医疗保健可解释预测模型》, 《神经信息处理系统进展》第29卷（2016年）。

[22] G. E. Box、G. M. Jenkins、G. C. Reinsel、G. M. Ljung，《时间序列分析：预测与控制》，John Wiley & Sons，2015年。

[23] T. L. Nguyen、S. Gsponer、I. Ilie、G. Ifrim，利用全子序列学习和时间域与频域中的符号表示实现可解释的时间序列分类，arXiv预印本 arXiv:1808.04022（2018年）。

[24] J. Wang、Z. Peng、X. Wang、C. Li、J. Wu，用于可解释多变量时间序列预测的深度模糊认知图，《IEEE模糊系统汇刊》29（2020年）2647–2660。

[25] G. Li、B. Choi、J. Xu、S. S. Bhowmick、K.-P. Chun、G. L.-H. Wong，ShapeNet：一种用于多变量时间序列分类的形状片段神经网络方法，载于：《AAAI人工智能会议论文集》，第35卷，2021年，第8375–8383页。

[26] Z. Zhao, Y. Shi, S. Wu, F. Yang, W. Song, N. Liu，《时间序列深度模型解读：综述》，arXiv预印本 arXiv:2305.14582（2023年）。

[27] Z. Sun, S.-B. Lin，《相关样本下的分布式学习》，IEEE信息论汇刊 68（2022年）6003–6020。

[28] M. T. Ribeiro, S. Singh, C. Guestrin，《“我为什么应该相信你？”：解释任意分类器的预测》，载于：第22届ACM SIGKDD知识发现与数据挖掘国际会议论文集，2016年，第1135–1144页。

[29] S. M. Lundberg, S.-I. Lee，《解释模型预测的统一方法》，神经信息处理系统进展30（2017年）。

[30] B. Mohammadi、N. Malik、T. Derdenger、K. Srinivasan，《监管可解释人工智能（XAI）可能会损害消费者利益》，《营销科学》44（2025年）711–724。

[31] M. Mueller-Peltzer、S. Feuerriegel、A. M. Nielsen、A. Kongsted、W. Vach、D. Neumann，《用于疾病管理的纵向医疗保健分析：腰痛的实证演示》，《决策支持系统》132（2020年）113271。

[32] Z. Shao, X. Yao, F. Chen, Z. Wang, J. Gao，重新审视股市预测中的时变动态：一种基于大语言模型的多源情感分析方法，《决策支持系统》190 (2025年) 114362。

[33] U. A. Victor-Edema, I. D. Essi，带外生变量的自回归积分滑动平均（ARIMAX）模型在尼日利亚非石油出口中的应用，《欧洲商业与管理杂志》8 (2016年) 29–34。

[34] J. H. Stock, M. W. Watson，向量自回归，《Journal ofEconomic Perspectives》15 (2001年) 101–115。

[35] T. Sivill, P. Flach，Limesegment：有意义且贴近现实的时间序列解释，载于：国际人工智能与统计会议，PMLR，2022年，第3418–3433页。

[36] J. Eun Choi、J. Won Shin、D. Wan Shin，机器学习时间序列预测的向量 SHAP 值，《预测学杂志》44（2025年）635–645。

[37] J. Crabb´e、M. Van Der Schaar，使用动态掩码解释时间序列预测结果，载于：国际机器学习会议，PMLR，2021年，第2166–2177页。

[38] S. M. Jankovi´c、A. V. Pej˘ci´c、M. N. Milosavljevi´c、V. D. Opan˘cina、N. V. Pe˘si´c、T. T. Nedeljkovi´c、G. M. Babi´c，重症监护病房患者潜在药物相互作用的风险因素，《重症监护杂志》43（2018年）1–6。

[39] C.-P. Tsai、C.-K. Yeh、P. Ravikumar，Faith-SHAP：忠实的 Shapley 交互指数，《机器学习研究杂志》24（2023年），1–42。

[40] M. Sundararajan、K. Dhamdhere、A. Agarwal，Shapley Taylor 交互指数，载于：国际机器学习会议，PMLR，2020年，第9259–9268页。

[41] C.-H. Chang、R. Caruana、A. Goldenberg，NODE-GAM：用于可解释深度学习的神经广义加性模型，arXiv预印本 arXiv:2106.01613（2021年）。

[42] J. D. Janizek、P. Sturmfels、S.-I. Lee，《解释解释：深度网络的公理化特征交互》，《机器学习研究杂志》22（2021年），1–54。

[43] H. Wendland，《散乱数据逼近》，第17卷，剑桥大学出版社，2004年。

[44] S. Lin、J. Zeng，多项式核快速学习，《IEEE 控制论汇刊》49（2018年），3780–3792。

[45] J. Zeng、M. Wu、S.-B. Lin、D.-X. Zhou，面向海量数据的快速多项式核分类，arXiv预印本 arXiv:1911.10558（2019年）。

[46] R. F. Bass、K. Gr¨ochenig，《多元三角多项式的随机采样》，SIAM《数学分析杂志》36 (2005) 773–795。

[47] Y. Guo、Y. Li、D. Liu、S. X. Xu，《基于顾客情绪的服务质量测量：一种可解释人工智能方法》，《决策支持系统》176 (2024年) 114051。