

# 基于LightGBM的多步时间序列预测架构深度研究与策略优化报告

在现代数据科学与预测分析领域，时间序列预测已成为决策支持系统的核心组成部分，广泛应用于金融市场分析、能源需求预测、供应链优化及宏观经济指标研判等关键领域<sup>1</sup>。随着机器学习技术的演进，传统的统计学模型如自回归移动平均模型(ARIMA)在处理海量、非线性、高维特征数据时逐渐显露局限，而以LightGBM(Light Gradient Boosting Machine)为代表的梯度提升决策树(GBDT)框架，凭借其卓越的计算效率、内存优化以及处理非平稳数据的能力，已成为工业界处理时间序列任务的首选工具<sup>3</sup>。

多步时间序列预测旨在根据历史观测序列  $y_1, y_2, \dots, y_t$  预测未来  $H$  个时间点的数值  $y_{t+1}, y_{t+2}, \dots, y_{t+H}$ ，其中  $H$  被称为预测界限或预测视野(Forecast Horizon)<sup>6</sup>。相较于单步预测，多步预测面临着累积误差、时间相依性减弱以及特征可用性受限等复杂挑战<sup>8</sup>。本报告旨在深度剖析基于LightGBM的多步预测四大核心策略——递归预测(Recursive)、直接预测(Direct)、直接递归混合预测(DirRec)以及多输出预测(Multi-output)，并针对实际开发中遇到的架构缺陷提供系统性的完善方案与理论指导。

## 多步预测策略的理论演进与分类架构

多步预测策略的选择直接决定了模型在不同预测深度下的准确性与稳健性。根据模型对预测界限的处理方式以及预测值之间的相互依赖关系，预测方法可归纳为单输出迭代法、独立并行法及其衍生的链式与联合优化法<sup>8</sup>。

### 递归预测(Recursive Strategy)的内在机制与局限

递归策略，亦被称为迭代预测(Iterated Forecasting)，其核心思想是训练一个单一的基模型  $f$  用于单步预测，即  $y_{t+1} = f(y_t, y_{t-1}, \dots, y_{t-k})$ <sup>8</sup>。在推理阶段，模型首先生成  $t + 1$  时刻的预测值  $\hat{y}_{t+1}$ ，随后将该预测值作为已知的观测数据反馈至特征矩阵中，用于生成  $t + 2$  时刻的预测，依此类推直至完成  $H$  步预测<sup>7</sup>。

从计算效率角度看，递归法极具优势，因为在整个预测生命周期中仅需维护一个模型，这显著降低了训练成本和内存占用<sup>9</sup>。然而，这种方法存在严重的“误差传播”(Error Propagation)缺陷<sup>8</sup>。由于  $\hat{y}_{t+1}$  并非真实观测值，其包含的预测偏误会在迭代过程中不断累积并放大，导致预测曲线在远端往往收敛至训练集的均值或出现不真实的波动<sup>9</sup>。

### 直接预测(Direct Strategy)的并行架构

直接预测策略采取了完全不同的路径，它不依赖于之前的预测值，而是为预测视野中的每一个步骤  $h \in \{1, \dots, H\}$  分别训练一个独立的模型  $f_h$ <sup>6</sup>。每个模型  $f_h$  的目标函数定义为  $y_{t+h} = f_h(y_t, y_{t-1}, \dots, y_{t-k})$ ，这意味着在特征工程阶段，所有的输入特征必须仅基于截止到  $t$  时刻的已知信息<sup>11</sup>。

直接法的最大优势在于它从根源上消除了递归法中的误差传播问题，因为每个预测步都是基于原始历史数据直接产生的<sup>8</sup>。然而，该策略的成本在于极高的计算开销，尤其是当预测视野  $H$  较大时，需要训练和存储海量的模型文件<sup>6</sup>。此外，由于各预测步的模型互不通信，直接法往往忽略了未来值之间的时序相关性(Conditional Dependence)，导致生成的预测序列在平滑度和逻辑一致性上表现欠佳<sup>8</sup>。

### 混合策略与联合优化 : DirRec 与 Multi-output

为了平衡递归法的效率与直接法的稳健性，DirRec(Direct-Recursive)策略通过“链式增强”的方式整合了二者的优点<sup>7</sup>。在 DirRec 架构中，为每一步训练独立模型(类似直接法)，但在训练第  $h$  个模型时，输入特征不仅包含历史数据，还包含前  $h - 1$  个步骤的预测值<sup>6</sup>。这种方法通过 RegressorChain 实现，能够允许后续模型对前序模型的预测偏差进行“修正”<sup>19</sup>。

多输出策略(Multi-output Strategy)则试图在单一优化框架内同时解决  $H$  个目标的预测问题<sup>21</sup>。虽然 LightGBM 原生并不支持多目标回归，但通过 MultiOutputRegressor 包装器，可以实现多目标损耗函数的并行拟合<sup>18</sup>。

策略维度	递归预测 (Recursive)	直接预测 (Direct)	直接递归预测 (DirRec)	多输出预测 (Multi-output)
模型数量	1	$H$	$H$	$H$ (封装后)
训练复杂度	低	高	极高	中
误差累积	严重	无	中度	无
时序依赖捕捉	隐含在迭代中	弱	强 (链式)	强 (联合)
应用场景	短期/计算资源	长期/准确性要	复杂相依性序	静态特征/多目

	受限	求高	列	标预测
--	----	----	---	-----

6

## LightGBM 核心算法及其在时序任务中的优势

LightGBM 之所以能在大规模时间序列任务中击败传统的统计模型及部分神经网络，归功于其底层设计的几大创新，这些设计直接针对了时序特征工程中常见的高维稀疏数据及计算瓶颈<sup>5</sup>。

### 基于直方图的决策树优化

传统的预排序(Pre-sort)算法在处理大规模时序数据时，需遍历所有可能的分割点，计算代价昂贵<sup>26</sup>。LightGBM 引入了基于直方图的算法(Histogram-based algorithm)，将连续的特征值分桶(Binning)至离散的直方图中<sup>9</sup>。这一改进将计算复杂度从  $O(\#data)$  降低至  $O(\#bins)$ ，不仅大幅提升了训练速度，还显著减少了内存消耗，使得在处理具有数千个滞后特征的特征矩阵时依然游刃有余<sup>26</sup>。

### 深度优先的 Leaf-wise 生长策略

与 XGBoost 等框架采用的 Level-wise(层优先)生长策略不同，LightGBM 采用 Leaf-wise(叶子优先)生长策略<sup>9</sup>。该算法每次分裂具有最大增益(Gradient Gain)的节点，从而能够以更浅的树深度达到更高的预测精度<sup>25</sup>。对于时序预测而言，这种策略能更敏锐地捕捉到那些细微但关键的周期性波动，但也伴随着更高的过拟合风险<sup>25</sup>。因此，在配置 num\_leaves 和 min\_data\_in\_leaf 时需格外谨慎，以防止模型陷入局部噪声的拟合中<sup>27</sup>。

### GOSS 与 EFB：应对特征爆炸的利器

在进行多步预测时，研究者通常会构造复杂的特征集，包括不同频率的滞后特征(Lag Features)和滚动窗口统计量(Rolling Statistics)<sup>3</sup>。这导致特征空间维度激增。

- **GOSS(单边梯度采样)**：通过保留梯度较大的样本并对梯度较小的样本进行随机采样，GOSS 在保证模型精度的同时，减少了计算分裂收益所需的数据量<sup>5</sup>。
- **EFB(互斥特征捆绑)**：时序特征中常存在大量互斥特征(如独热编码后的类别特征)。EFB 将这些特征捆绑在一起，有效降低了特征维度，进一步优化了训练效率<sup>9</sup>。

## 时间序列特征工程的工程实践与深度洞察

在 LightGBM 的语境下，时间序列预测本质上被转化为监督学习任务。特征工程的质量直接决定了模型能否从历史数据中提取有效的时序模式<sup>10</sup>。

### 滞后特征(Lag Features)的构造逻辑

滞后特征是捕获序列自相关性(Autocorrelation)的基础。滞后  $k$  代表  $t - k$  时刻的观测值<sup>3</sup>。

在多步预测中，滞后特征的选取必须遵循“因果一致性”原则<sup>7</sup>。对于直接预测模型  $f_h$ ，其输入特征集  $\mathcal{X}_t$  仅能包含  $y_t, y_{t-1}, \dots$  及其衍生指标，绝对不能包含任何在预测发生点  $t$  之后的信息<sup>17</sup>。一种常见的策略是利用偏自相关函数(PACF)图来确定显著的滞后点，例如对于具有周季节性的数据，滞后 1、7、14、21 是至关重要的特征<sup>7</sup>。

## 动态窗口与滚动统计量

单一的时间点数据往往包含较多随机噪声，通过构造滚动窗口统计量(如滚动均值、滚动标准差、滚动最大/最小值)，可以赋予模型识别局部趋势和波动率变化的能力<sup>3</sup>。例如，rolling\_mean\_7 可以反映过去一周的基准需求，而 rolling\_std\_30 则能刻画过去一个月的市场稳定性<sup>10</sup>。这些特征在递归预测中需要随每一步的预测值动态重新计算，而在直接预测中则是基于  $t$  时刻之前的固定窗口计算<sup>3</sup>。

## 外生变量(Exogenous Variables)的集成与挑战

LightGBM 处理多源异构数据的能力使其在集成外生变量(如节假日、气象指标、经济指数)时表现优异<sup>2</sup>。

- 静态特征：如店铺 ID、地理位置，可通过 LightGBM 原生的 categorical\_feature 参数进行优化处理<sup>35</sup>。
- 未来已知外生变量：如未来的节假日标识或已排定的促销活动。这些变量在多步预测中极为珍贵，因为它们是“未来确定性”的来源<sup>17</sup>。在递归预测中，如果缺少未来的气象预测值，通常需要使用历史同期的平均值或采用专门的补偿机制进行填充，以维持预测链条的完整性<sup>38</sup>。

特征类别	描述与公式	在多步预测中的作用
自回归滞后	$y_{t-1}, y_{t-}$	捕捉短期动量与季节性周期
滚动均值	$\frac{1}{w} \sum_{i=1}^w y_{t-i}$	平滑噪声，识别局部趋势
时间特征	月份、周几、小时	编码确定性的周期性波动
节假日指示器	二进制开关 (0/1)	处理由特殊事件引起的非规律性脉冲
交互特征	滞后值 $\times$ 趋势因子	捕捉变量间的协同效应

## 四大预测方法的完美实现与架构优化方案

针对开发者在 DirRec 和 Multi-output 方法实现中遇到的技术瓶颈，本节提供了基于 LightGBM 与 Scikit-learn 的标准工业化实现逻辑。

### 递归法(**Recursive**)的规范化实现

递归法的核心在于预测循环与特征更新的同步。在每一轮循环中，模型必须重构输入特征行，确保最新的预测值被放置在“滞后 1”的位置。

Python

```
# 递归预测核心逻辑伪代码
def recursive_forecast(model, last_window, horizon):
    forecasts =
    current_window = last_window.copy()
    for h in range(horizon):
        # 1. 提取当前特征集(包含动态生成的滞后项)
        features = extract_features(current_window)
        # 2. 预测下一步
        pred = model.predict(features)
        forecasts.append(pred)
        # 3. 核心步骤: 更新滑动窗口, 将预测值注入
        current_window = update_window(current_window, pred)
        # 4. 重新计算依赖于新预测值的滚动特征
        current_window = recalculate_rolling_stats(current_window)
    return forecasts
```

12

### 直接法(**Direct**)的并行化实现

直接法要求特征集在所有步骤中保持一致，但目标变量随步骤平移。这种方法易于并行化，利用 n\_jobs 参数可加速训练<sup>6</sup>。

Python

```

# 直接预测封装逻辑
from sklearn.multioutput import MultiOutputRegressor
# 虽然名为 MultiOutput, 但内部机制是为每个 Target 拟合独立模型
direct_model = MultiOutputRegressor(lgb.LGBMRegressor(**params), n_jobs=-1)
direct_model.fit(X_train, Y_train_multi_step) # Y 为 (n_samples, horizon)

```

18

## DirRec(链式递归 ) 的修复与优化方案

DirRec 方法常出错的原因在于特征维度的动态增长。使用 RegressorChain 是解决此问题的标准方式, 它自动管理特征链条的传递<sup>8</sup>。

完善要点:

1. 目标排序: 目标矩阵  $Y$  的列顺序必须严格按照时间  $t + 1, t + 2, \dots$  排列<sup>19</sup>。
2. 特征一致性: 训练时的输入特征  $X$  仅包含  $t$  之前的信息, 链式模型会自动在推理阶段将  $\hat{y}_{t+1}$  拼接到  $X$  后作为  $f_2$  的输入<sup>20</sup>。

Python

```

from sklearn.multioutput import RegressorChain
# 修复方案: 使用 RegressorChain 并确保 order 逻辑正确
dirrec_model = RegressorChain(lgb.LGBMRegressor(**params), order=[0, 1, 2, ..., H-1])
dirrec_model.fit(X_train, Y_train_multi_step)

```

8

## 多输出 (Multi-output) 的联合优化策略

在 LightGBM 中实现多输出, 虽然本质上也是拟合多个模型, 但其优势在于封装性。修复多输出方法的关键在于正确对齐  $Y$  矩阵<sup>21</sup>。

完善要点:

1. 损失函数对齐: 对于多步预测, 通常使用全局 RMSE 作为评估指标, 确保所有步长的误差都被同等重视<sup>18</sup>。

2. 内存管理:对于极长的预测视野  $H$ , 需在预测后及时清理模型对象或使用批处理推理, 防止内存溢出<sup>18</sup>。

## 模型评估、验证与性能调优框架

时间序列预测的特殊性在于数据的时间依赖性。传统的交叉验证(K-Fold)会导致严重的“数据泄露”, 即利用未来信息预测过去<sup>3</sup>。

### 走动验证(Walk-Forward Validation)与回测

理想的验证方案是基于时间的切分(TimeSeriesSplit)或回测(Backtesting)<sup>10</sup>。

- 不重拟合回测:模型训练一次, 在测试集上以固定步长滑动预测, 用于评估模型在真实生产环境中的长期衰减情况<sup>40</sup>。
- 带重拟合回测:每次滑动后根据新收到的观测值重新训练模型, 这能评估模型对最新趋势的捕捉能力, 但计算成本极高<sup>34</sup>。

### 关键超参数的调优艺术

LightGBM 在多步预测任务中的表现高度依赖于参数微调。针对时序任务的典型配置如下:

参数名称	建议范围	调优逻辑
num_leaves	15 - 64	控制树复杂度; 时序数据复杂时可适当增加, 但需配合正则化 <sup>13</sup>
learning_rate	0.01 - 0.1	较小的学习率配合更多的n_estimators通常能获得更稳健的收敛 <sup>13</sup>
feature_fraction	0.7 - 0.9	每次分裂考虑的特征比例, 有助于防止某几个强滞后特征过早主导树结构 <sup>13</sup>
lambda_l1 / l2	0.0 - 10.0	增加正则化项以应对多步预测中特征高度共线性的问题 <sup>29</sup>
min_data_in_leaf	20 - 500	防止叶子节点过小导致的局部噪声拟合 <sup>27</sup>

## 全球模型与多序列预测：规模化的范式

在处理诸如零售链中数千种商品的销量预测时，为每个序列建立独立模型已变得不切实际。目前的最佳实践是构建“全球模型”(Global Models)<sup>12</sup>。

### 跨序列信息共享

全球模型将所有的单变量序列拼接成一个巨大的数据集。通过引入序列 ID 作为分类特征，LightGBM 可以学习到跨序列的共同模式<sup>12</sup>。例如，在节假日期间，所有商品的销量可能都会激增，全球模型能从全局数据中学习到这种一致性的脉冲响应，即使某个特定商品的短历史数据不足以推断出这一规律<sup>12</sup>。

### 独立多序列递归预测

利用 ForecasterRecursiveMultiSeries(如 skforecast 或 mlforecast 中的实现)，研究者可以轻松管理数千个序列的递归流程<sup>34</sup>。这种架构不仅提升了模型的稳健性，还极大地简化了维护流程，因为仅需监控一个全局模型的性能指标<sup>34</sup>。

## 结论与专家建议

综上所述，基于 LightGBM 的多步时间序列预测并非单一的方法选择，而是根据业务场景、数据规模及预测视野深度进行的多维权衡。

1. 对于短中期预测：建议优先采用递归预测方法，通过精细的特征工程和严格的正则化来控制误差传播，其优势在于能够充分利用最新的预测反馈<sup>9</sup>。
2. 对于长期准确性要求极高的场景：直接预测策略虽然计算成本高，但其对误差累积的免疫力使其成为更安全的选择<sup>6</sup>。
3. 解决 DirRec 与 Multi-output 的技术难题：应标准化使用 RegressorChain 和 MultiOutputRegressor 封装器，并确保目标矩阵  $Y$  的时序对齐逻辑严密无偏<sup>8</sup>。
4. 特征驱动核心：无论选择何种预测方法，构造高质量的滞后特征、动态滚动统计量以及集成确定性的外生变量始终是提升预测精度的关键所在<sup>10</sup>。

未来，随着 Text-Guided 等跨模态预测技术的成熟<sup>42</sup>，将非结构化文本信息与 LightGBM 的高效决策机制相结合，有望在剧烈波动的市场环境下进一步推高时间序列预测的准确率上限。

### 引用的著作

1. Time Series Forecasting with PyCaret: Building Multi-Step Prediction Model - MachineLearningMastery.com, 访问时间为二月 3, 2026,  
<https://machinelearningmastery.com/time-series-forecasting-with-pycaret-building-multi-step-prediction-model/>
2. What Is Multivariate Time Series Forecasting With Exogenous Variables? - Learn

About Economics - YouTube, 访问时间为 二月 3, 2026,

<https://www.youtube.com/watch?v=oZdRPlx2kYE>

3. Time Series using LightGBM - GeeksforGeeks, 访问时间为 二月 3, 2026,  
<https://www.geeksforgeeks.org/machine-learning/time-series-using-lightgbm/>
4. Multi-step Time Series Forecasting with ARIMA, LightGBM, and Prophet, 访问时间为 二月 3, 2026,  
<https://towardsdatascience.com/multi-step-time-series-forecasting-with-arima-lightgbm-and-prophet-cc9e3f95dfb0/>
5. Regression using LightGBM - GeeksforGeeks, 访问时间为 二月 3, 2026,  
<https://www.geeksforgeeks.org/machine-learning/regression-using-lightgbm/>
6. Multistep Forecasting Strategies - Kaggle, 访问时间为 二月 3, 2026,  
<https://www.kaggle.com/code/eugeniyosetrov/multistep-forecasting-strategies>
7. Forecasting in Python: A Practical Guide | by Alan Vourc'h - Medium, 访问时间为 二月 3, 2026,  
<https://medium.com/@alan.vourch/forecasting-a-practical-guide-6173f421c1ed>
8. 6 Methods for Multi-step Forecasting | Towards Data Science, 访问时间为 二月 3, 2026,  
<https://towardsdatascience.com/6-methods-for-multi-step-forecasting-823cbde4127a/>
9. LightGBM for TimeSeries forecasting | by Michele Pace | Data Reply IT | DataTech - Medium, 访问时间为 二月 3, 2026,  
<https://medium.com/data-reply-it-datatech/lightgbm-for-timeseries-forecasting-408971289a12>
10. Time Series Forecasting - Abilian Innovation Lab, 访问时间为 二月 3, 2026,  
<https://lab.abilian.com/Tech/AI%20%26%20Machine%20Learning/Time%20Series%20Forecasting/>
11. Intro to Forecasting - Skforecast Docs, 访问时间为 二月 3, 2026,  
<https://skforecast.org/0.18.0/introduction-forecasting/introduction-forecasting>
12. Recursive MultiStep Time Series Forecasting - Kaggle, 访问时间为 二月 3, 2026,  
<https://www.kaggle.com/code/ahmedabdulhamid/recursive-multistep-time-series-forecasting>
13. Time Series Forecasting with LightGBM | by Rani RATNA SRI - Medium, 访问时间为 二月 3, 2026,  
<https://medium.com/@raniratnasri/time-series-forecasting-with-lightgbm-39c9e6c93f87>
14. ML time series forecasting the right way | by Mario Dagrada | TDS Archive - Medium, 访问时间为 二月 3, 2026,  
<https://medium.com/data-science/ml-time-series-forecasting-the-right-way-cbf3678845ff>
15. Direct forecast X Recursive forecast - R-bloggers, 访问时间为 二月 3, 2026,  
<https://www.r-bloggers.com/2018/01/direct-forecast-x-recursive-forecast/>
16. Direct multi-step forecasting - Skforecast Docs, 访问时间为 二月 3, 2026,  
[https://skforecast.org/0.16.0/user\\_guides/direct-multi-step-forecasting.html](https://skforecast.org/0.16.0/user_guides/direct-multi-step-forecasting.html)
17. Exogenous features with global models · Issue #432 · Nixtla/mlforecast - GitHub, 访问时间为 二月 3, 2026, <https://github.com/Nixtla/mlforecast/issues/432>

18. Multi-Target Sales Forecasting LGBM - Kaggle, 访问时间为 二月 3, 2026,  
<https://www.kaggle.com/code/stpeteishii/multi-target-sales-forecasting-lgbm>
19. RegressorChain — scikit-learn 1.8.0 documentation, 访问时间为 二月 3, 2026,  
<https://scikit-learn.org/stable/modules/generated/sklearn.multioutput.RegressorChain.html>
20. Chained multi-output regression solution with Scikit-Learn | Towards ..., 访问时间为 二月 3, 2026,  
<https://towardsdatascience.com/chained-multi-output-regression-solution-with-scikit-learn-4f44bf9c8c5b/>
21. Multioutput Regression in Machine Learning - GeeksforGeeks, 访问时间为 二月 3, 2026,  
<https://www.geeksforgeeks.org/machine-learning/multioutput-regression-in-machine-learning/>
22. How to Develop Multi-Output Regression Models with Python - Machine Learning Mastery, 访问时间为 二月 3, 2026,  
<https://machinelearningmastery.com/multi-output-regression-models-with-python/>
23. MultiOutputRegressor — scikit-learn 1.8.0 documentation, 访问时间为 二月 3, 2026,  
<https://scikit-learn.org/stable/modules/generated/sklearn.multioutput.MultiOutputRegressor.html>
24. regression with scikit-learn with multiple outputs, svr or gbm possible? - Cross Validated, 访问时间为 二月 3, 2026,  
<https://stats.stackexchange.com/questions/153853/regression-with-scikit-learn-with-multiple-outputs-svr-or-gbm-possible>
25. XGBoost vs LightGBM - It's Amit - Medium, 访问时间为 二月 3, 2026,  
<https://mr-amit.medium.com/xgboost-vs-lightgbm-b6ca76620156>
26. Features — LightGBM 4.6.0.99 documentation, 访问时间为 二月 3, 2026,  
<https://lightgbm.readthedocs.io/en/latest/Features.html>
27. XGBoost vs LightGBM: How Are They Different - Neptune.ai, 访问时间为 二月 3, 2026, <https://neptune.ai/blog/xgboost-vs-lightgbm>
28. RecLGB: Enhancing LightGBM using Recursive VAE with Mixed Attention for Time-Series Forecasting - ResearchGate, 访问时间为 二月 3, 2026,  
[https://www.researchgate.net/publication/397633766\\_RecLGB\\_Enhancing\\_LightGBM\\_using\\_Recursive\\_VAE\\_with\\_Mixed\\_Attention\\_for\\_Time-Series\\_Forecasting](https://www.researchgate.net/publication/397633766_RecLGB_Enhancing_LightGBM_using_Recursive_VAE_with_Mixed_Attention_for_Time-Series_Forecasting)
29. LightGBM - Parameter Tuning - TutorialsPoint, 访问时间为 二月 3, 2026,  
<https://www.tutorialspoint.com/lightgbm/lightgbm-parameter-tuning.htm>
30. Lagged features for time series forecasting - Scikit-learn, 访问时间为 二月 3, 2026  
,  
[https://scikit-learn.org/stable/auto\\_examples/applications/plot\\_time\\_series\\_lagged\\_features.html](https://scikit-learn.org/stable/auto_examples/applications/plot_time_series_lagged_features.html)
31. Introduction to forecasting - Skforecast Docs, 访问时间为 二月 3, 2026,  
<https://skforecast.org/0.12.1/introduction-forecasting/introduction-forecasting>
32. Future Looking Target Variables in Light GBM - Stack Overflow, 访问时间为 二月 3, 2026,

<https://stackoverflow.com/questions/79657687/future-looking-target-variables-in-light-gbm>

33. Feature selection strategies for multivariate time series forecasting :  
r/MLQuestions - Reddit, 访问时间为 二月 3, 2026,  
[https://www.reddit.com/r/MLQuestions/comments/1q0a3lj/feature\\_selection\\_strategies\\_for\\_multivariate/](https://www.reddit.com/r/MLQuestions/comments/1q0a3lj/feature_selection_strategies_for_multivariate/)
34. Global Models : Independent multi-time series forecasting - Skforecast Docs, 访问时间为 二月 3, 2026,  
[https://skforecast.org/0.13.0/user\\_guides/independent-multi-time-series-forecasting](https://skforecast.org/0.13.0/user_guides/independent-multi-time-series-forecasting.html)
35. Top 7% Score: Simple LightGBM Timeseries Forecast - Kaggle, 访问时间为 二月 3, 2026,  
<https://www.kaggle.com/code/tommadden/top-7-score-simple-lightgbm-timeseries-forecast>
36. Comparative Analysis of Global and Local Probabilistic Time Series Forecasting for Contiguous Spatial Demand Regions - arXiv, 访问时间为 二月 3, 2026,  
<https://arxiv.org/html/2509.08214>
37. Handling Categorical Features using LightGBM - GeeksforGeeks, 访问时间为 二月 3, 2026,  
<https://www.geeksforgeeks.org/machine-learning/handling-categorical-features-using-lightgbm/>
38. Forecasting with Exogenous Regressors - arch 7.2.0, 访问时间为 二月 3, 2026,  
[https://arch.readthedocs.io/en/latest/univariate/univariate\\_forecasting\\_with\\_exogenous\\_variables.html](https://arch.readthedocs.io/en/latest/univariate/univariate_forecasting_with_exogenous_variables.html)
39. LightGBM Medium-Term Photovoltaic Power Prediction Integrating Meteorological Features and Historical Data - MDPI, 访问时间为 二月 3, 2026,  
<https://www.mdpi.com/1996-1073/18/20/5526>
40. Intro to Forecasting - Skforecast Docs, 访问时间为 二月 3, 2026,  
<https://skforecast.org/0.19.1/introduction-forecasting/introduction-forecasting.html>
41. Independent multi-time series forecasting - Skforecast Docs, 访问时间为 二月 3, 2026,  
[https://skforecast.org/0.19.1/user\\_guides/independent-multi-time-series-forecasting](https://skforecast.org/0.19.1/user_guides/independent-multi-time-series-forecasting.html)
42. Beyond Trend and Periodicity: Guide Time Series Forecasting with Textual Cues, 访问时间为 二月 3, 2026, <https://openreview.net/forum?id=mfc6FKgtQA>