

时序预测系列（一） 基于分解算法和深度学习的预测建模研究

投资咨询业务资格：
证监许可【2012】669号

报告要点

本次报告将分解算法（EMD 和 CEEMDAN）和深度学习算法（LSTM 和 GRU）分别组合对日度收盘价进行单步预测，发现价格经分解算法分解重组后可提升深度学习时序预测的有效性，若将分解后的高频序列进行去噪可以进一步强化预测模型的有效性。

摘要：

分解算法（EMD 和 CEEMDAN）能根据信号特点自适应地将信号分解成一组具有物理意义的 IMF 分量的线性组合，非常适合非线性、非平稳信号分析。深度学习算法（LSTM 和 GRU）对复杂序列中长期依赖关系有较强的提取能力。因此，本文将两者结合起来探究高精度时序预测的可能性。

为测试模型的有效性和泛化能力，本文在 CU、IF 和 T 三种期货合约上进行了测试。在三个品种的预测中， R^2 最高可提升 10%以上，三个品种的最佳方向准确率均达到 60%以上，我们发现在分解算法将价格序列分解成 IMF 分量组合后，可以不简单依靠 T 检验对序列进行高频、低频和趋势项的分类加和，将 IMF 分量依次组合求得最优组合方式能强化深度学习模型的学习能力。同时，通过剔除少量高频 IMF 分量对高频序列进行去噪可以再次增强优化效果。

此外，我们还发现虽然 EMD 和 CEEMDAN 均能对深度学习算法进行优化，但优化效果不一定，CEEMDAN 的优化效果不一定强于 EMD。同时，重组去噪后的 GRU 模型并非绝对优于重组去噪后的 LSTM 模型。

商品量化组

研究员：
蒋可欣 FRM
jiangkexin@citicsf.com
从业资格号 F03098078
投资咨询号 Z0018262

风险提示：本报告中所涉及的算法和模型应用仅为回溯举例，并不构成推荐建议。

目录

摘要:	1
一、引言	4
二、EMD 和 CEEMDAN 的基本原理和方法	5
(一) EMD	5
(二) CEEMDAN	7
三、LSTM 和 GRU 网络的内部结构和工作原理	8
(一) LSTM 神经网络	8
(二) GRU 神经网络	12
四、组合模型构建	14
(一) IMF 重组	14
(二) 组合模型建模流程	14
(三) LSTM 和 GRU 网络的计算图结构和训练方法选择	16
五、预测结果分析	17
(一) 数据选择和处理	17
(二) 预测结果评价指标	18
(三) 预测结果展示和分析	19
六、结论与展望	25
参考文献	26
免责声明	27

图表目录

图表 1: CU 主力合约价格走势和 EMD 分解过程	6
图表 2 : CEEMDAN 的更迭流程	7
图表 3: CU 主力合约价格走势和 CEEMDAN 分解过程	8
图表 4: 单个 RNN 神经元结构和展开形式	9
图表 5: LSTM 神经元结构	9
图表 6: Sigmoid 函数图像	10
图表 7: Tanh 函数图像	11
图表 8: GRU 神经元结构	12
图表 9: 组合模型的流程	15
图表 10: LSTM/GRU 网络的计算图结构	16
图表 11: CU 分解后各频率序列 LSTM 模型预测效果比较	19
图表 12: CU 经 EMD 分解高频序列模型预测结果展示	19
图表 13: CU 经 EMD 分解后高频序列模型预测效果比较	20
图表 14: CU 经 CEEMDAN 分解各序列 LSTM 模型预测效果比较	20
图表 15: CU 经 CEEMDAN 分解高频序列模型预测结果展示	21
图表 16: CU 各类模型预测结果展示	21

图表 17: CU 各类模型预测效果比较.....	22
图表 18: IF 各类模型预测结果展示.....	23
图表 19: IF 各类模型预测效果比较.....	23
图表 20: T 模型各类预测结果展示.....	24
图表 21: T 模型各类预测效果比较.....	24

一、引言

中国金融市场是一个受多种社会因素驱动的非线性复杂系统，价格时间序列通常具有显著非平稳、非线性和高噪音的复杂特性。当前市场，时序的预测模型主要有三类分别为 1) **技术分析法以道氏理论为基础**，认为股价基本走势与市场波动趋势趋同；2) **统计建模方法**依据严谨的统计学理论对价格时序进行预测建模；3) **利用机器学习方法**对金融时序进行预测建模。技术分析法比较直观，计算简单，但其信号延迟严重，所产生的买卖信号不确定性过高，在短期预测中效果不佳；统计建模方法的预测结果在统计意义上可靠，但其优势体现在处理线性数据上，需要严格的基础假设，实操性较弱；机器学习方法不需要满足统计假设，对非线性时间序列有着更高的预测精度。目前常用的机器学习方法包括决策树、支持向量机（SVM）和神经网络等，其中，神经网络的学习机制使其拥有强大的学习能力和非线性拟合能力，因此被广泛运用到价格时序预测中。近年来，传统的神经网络过拟合和泛化能力弱的问题被越来越多的学者验证，随着大数据技术的提高，深度学习技术受到了国内外学者的广泛关注。和传统的神经网络相比，深度学习模型层数更多，特征提取能力更强。然而，金融市场是一个受多重因素影响的复杂系统，随机性过多，即便使用了循环神经网络（RNN）和卷积神经网络（CNN）等为代表的深度学习技术，预测精度仍未得到充分开发。单个技术难以充分挖掘并刻画复杂金融市场中的多维量价变化规律，为克服依靠单一方法建模的局限性，融合计量手段、信号处理和机器学习的组合模型因其稳定的预测性能得到了广泛关注。因此，本次时序预测系列研究从**信号分析和深度学习**这两块展开，先通过各类信号分析方法对时序进行分解，再运用深度学习对分解后的分量分别进行预测。本篇为本系列的第一个课题，将对融合了分解算法和深度学习的组合模型进行讨论，分为四大部分，第一部分将介绍经验模态分解（EMD）及其后续优化算法自适应噪音的完整集成经验模态分解（CEEMDAN）的基本原理和方法；第二部分将介绍 LSTM 神经网络和 GRU 神经网络的内部结构和工作原理；第三部分是针对各分解算法和深度学习的复合模型构建；第四部分是对各复合模型预测效果的对比分析。此外，本系列虽以金融和商品期货为回测标的，但模型使用范围不限于期货，也可用于股票及其他适合的金融产品。

二、EMD 和 CEEMDAN 的基本原理和方法

(一) EMD

EMD 在 1998 年由美裔科学家 Huang 提出，它是一种数据驱动自适应时频分析方法。目前主要的时频分析方法有短时傅里叶变化和小波变换等，但它们都更适用于平稳序列，在实际使用时会经常失效。而 EMD 能根据信号特点自适应地将信号分解成一组具有物理意义的固有模态函数（IMF）的线性组合，且分解的最终函数彼此之间是正交的，因此比较适合非线性、非平稳信号分析。如公式 1 所示，信号 $s(t)$ 经过 EMD 可表示成从高频到低频的 IMF 分量的线性组合，其中 $IMF_i(t)$ 是第 i 个 IMF 分量， $r_N(t)$ 不包含信号的振荡模态，只是反映信号的变化趋势。

$$s(t) = \sum_{i=1}^N IMF_i(t) + r_N(t) \quad (\text{公式 1})$$

IMF 的计算过程主要分为 3 步：

1. 找出信号 $s(t)$ 的各个局部极大值，局部极大值的定义为时间序列中的某个时刻的值，其前一时刻的值不比它大，后一时刻的值也不比它大。然后用三阶样条函数进行插值，得到上包络曲线 $e_{max}(t)$ 。同理，可以得到下包络曲线 $e_{min}(t)$ 。

2. 对每个时刻的 $e_{max}(t)$ 和 $e_{min}(t)$ 取平均，得到瞬时平均值 $m(t)$ ：

$$m(t) = [e_{max}(t) + e_{min}(t)] / 2 \quad (\text{公式 2})$$

3. 用原信号 $s(t)$ 减去瞬时平均值 $m(t)$ ，得到类距平值系列：

$$d(t) = s(t) - m(t) \quad (\text{公式 3})$$

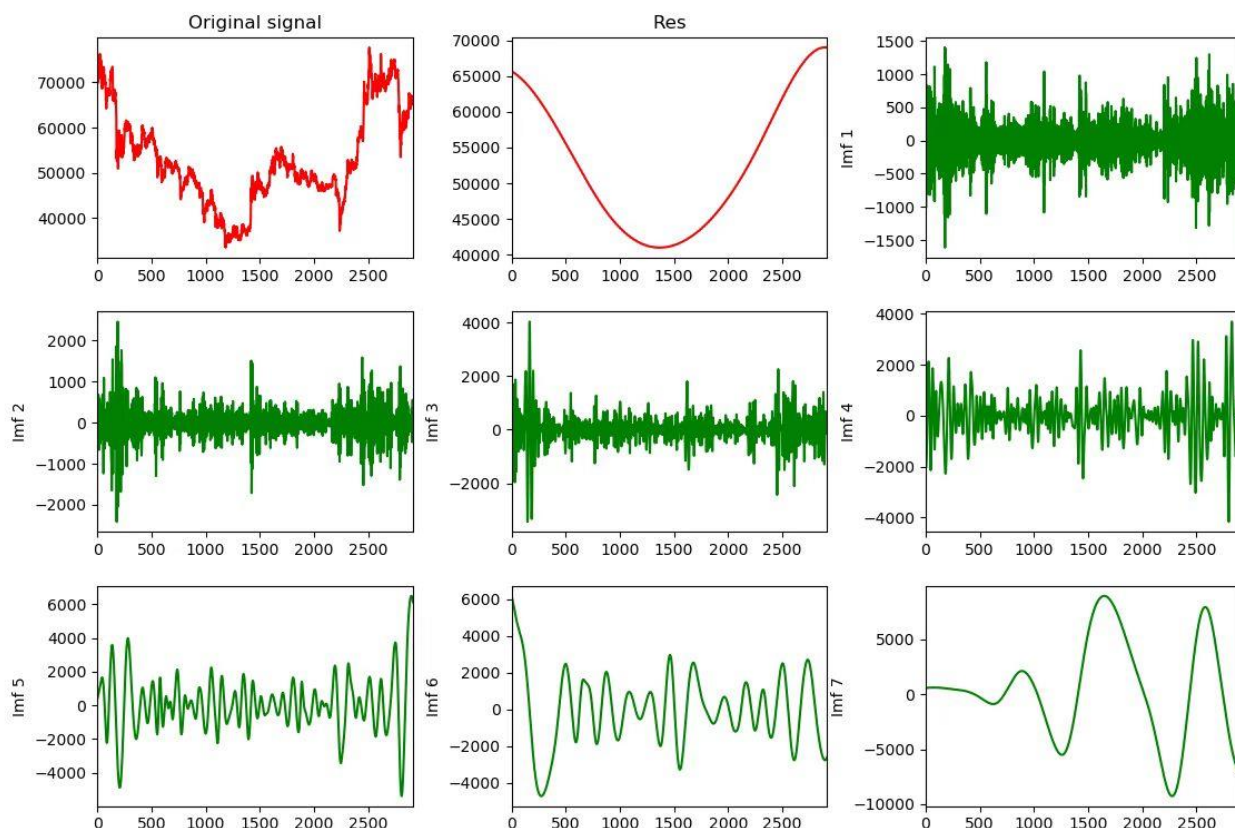
若 $d(t)$ 中极值点的数目和跨零点的数目相等或至多只差一个，并且各个瞬间平均值 $m(t)$ 都等于零，令 $d(t) = IMF_i(t)$ ，否则以 $d(t)$ 作为“原始”信号并返回步骤 1，重复以上步骤，直至满足 IMF 函数要求的要求，求出 IMF 函数为止。然后，用原信号减去 $IMF_i(t)$ ，得到剩余值序列 $r_i(t)$ ：

$$r_i(t) = s(t) - IMF_i(t) \quad (\text{公式 4})$$

若 $r_i(t)$ 是一个单调函数，把 $r_i(t)$ 作为趋势项，令 $r_i(t) = r_N(t)$ ，同时停止迭代，否则以 $r_i(t)$ 作为原信号，返回步骤 1，重复以上过程。这样一个流程就是

EMD 的分解流程，我们 CU 主力合约 2011 年 1 月 1 日至 2022 年 12 月 31 日的日频数据为例，直观地感受一下分解流程。图表 1 中的第一幅图为 IF 主力合约的价格走势，其余图对应了 EMD 分解过程，由图表 1 可知，该价格时序可以由 7 个固有模态函数和一个趋势项组成，原本振荡向下的价格时序最终被分解成一个单调向下的序列。

图表 1：CU 主力合约价格走势和 EMD 分解过程



资料来源：同花顺 iFind、中信期货研究所

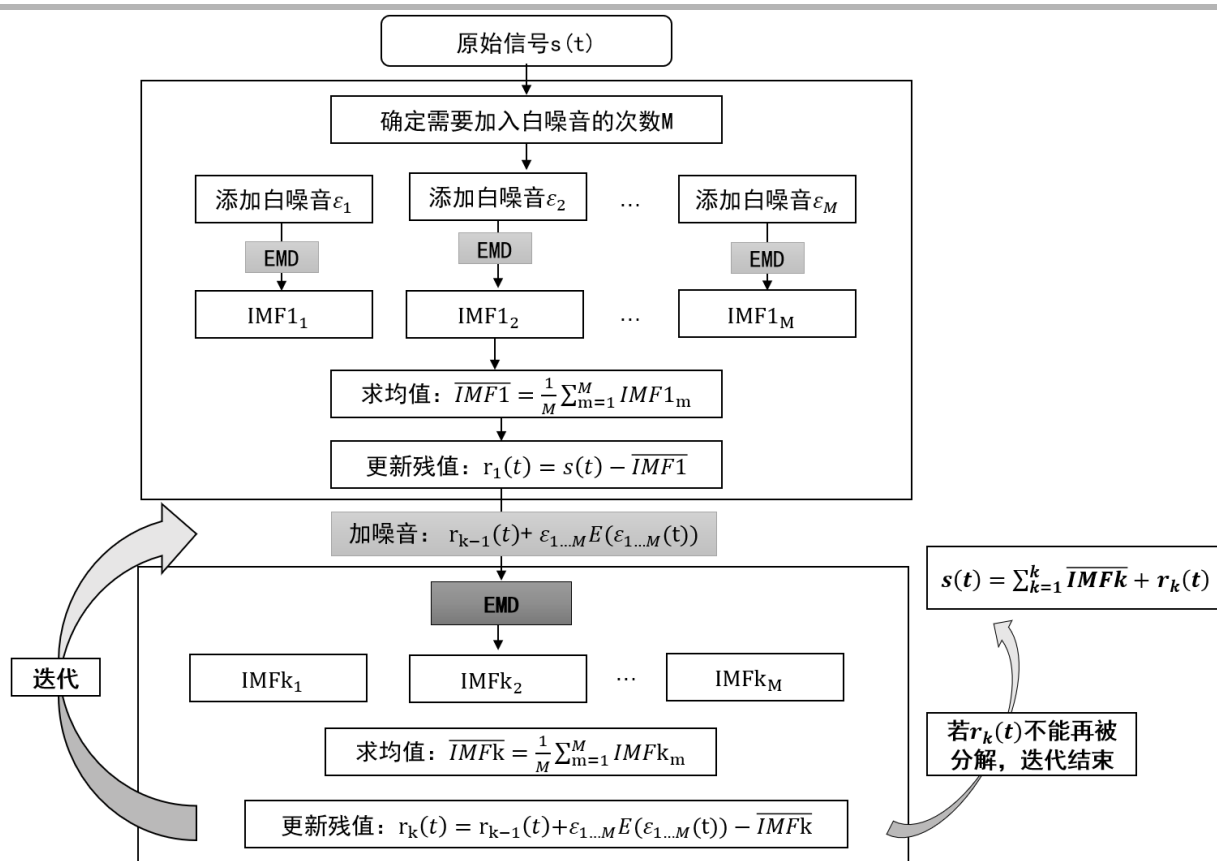
虽然 EMD 具有自适应性的特点且不用选择基函数，和传统的时频信号分析技术相比，分解非平稳非线性时序的优势明显，但存在严重的模态混叠现象，因此相关学者针对这个问题，提出了优化算法 EEMD。该方法通过在使用 EMD 方法之前在原序列中加入白噪音序列，以促进抗混分解，能有效的抑制混叠现象。那模态混叠是怎么产生的呢，我们在 EMD 分解过程中首先需要确定信号的局部极值点，然后用三次样条线将所有的局部极大值和极小值点分别连接起来形成上下包络曲线，再由上下包络曲线得到均值曲线。在求包络曲线的过程中，若信号中存在异常事件时，会影响极值点的选取，使极值点分布不均匀，从而导致求取的包络为异常事件的局部包络和真实信号包络的组合。经该包络计算出的均值，再筛选出的 IMF 分量就包含了信号的固有模式和异常事件或者包含了

相邻特征时间尺度的固有模式，从而产生了模态混叠现象。模态混叠现象具体表现为：1. 不同的特征频率出现在同一 IMF 分量中；2. 不同的 IMF 分量中存在相近的特征频率。价格时序中的异常事件通常有间断信号和脉冲干扰等，通过在初始信号中加入白噪音，可以平滑异常事件。但很快国内外学者发现 EEMD 对原序列添加的白噪音仍有可能在多次平均后影响分解产生的子序列，从而影响子序列的预测精度，为了抑制这个问题带来的影响，专家又提出了新一版的分解方法（CEEMDAN）。

（二）CEEMDAN

CEEMDAN 由 Torres 在 2011 年提出。EEMD 是直接原始信号上添加白噪音然后进行 EMD 分解，并对相对应的 IMF 直接求平均；而 CEEMDAN 不仅在原始信号上添加噪音，在每求完一阶 IMF 分量后，会重新给残值加入白噪音，具体更迭流程见图表 2。

图表 2：CEEMDAN 的更迭流程

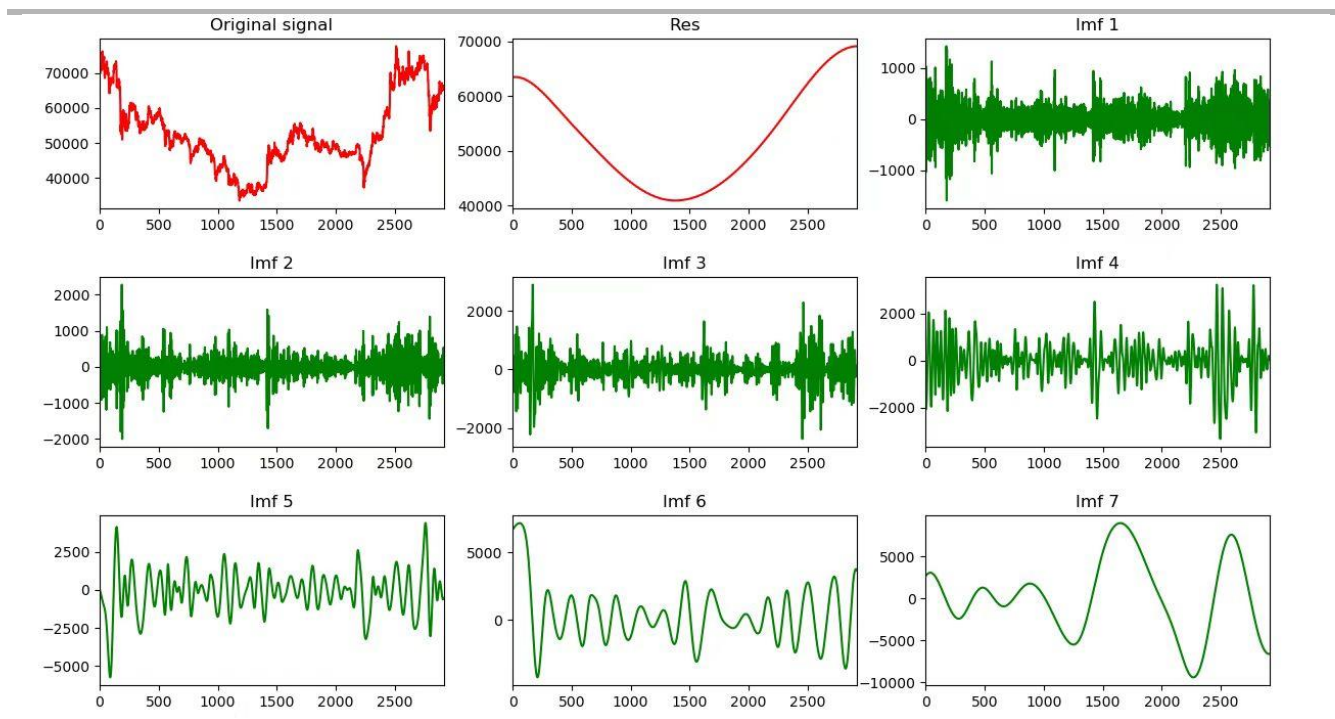


资料来源：中信期货研究所

从图表 2 中可以发现，CEEMDAN 在每一次迭代中都会对 IMF 进行求平均，已知白噪音的均值为零，这样可以有效地抑制白噪音从高频到低频的转移传递因

题。且 CEEMDA 是在每次分解后加入白噪音，它不用像 EEMD 在原信号上加很多次噪音，因此它的计算速度得到了提高。同样以 IF 主力合约的价格序列为例进行分解，图表 3 展示了 CEEMDAN 的分解过程。

图表 3：CU 主力合约价格走势和 CEEMDAN 分解过程



资料来源：同花顺 iFind、中信期货研究所

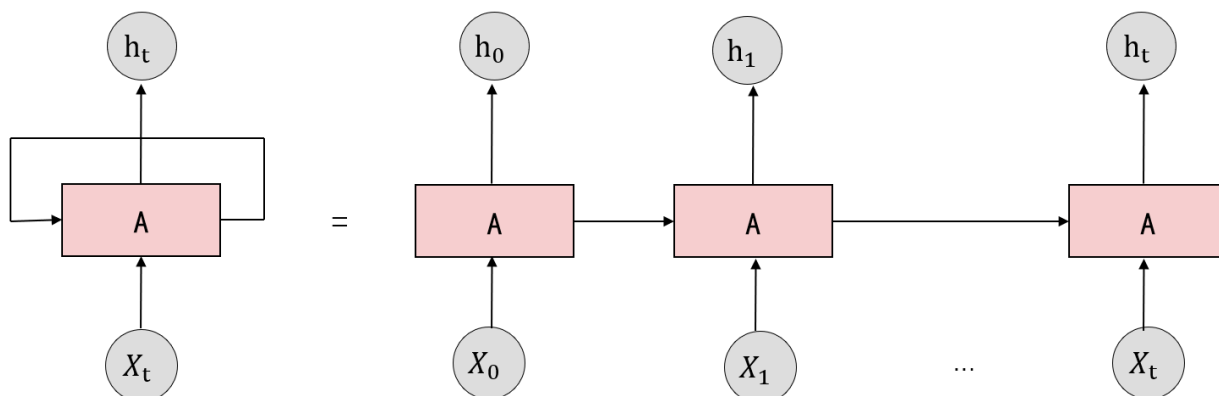
时序按 CEEMDAN 可分解成 5 个 IMF 函数和一个趋势项，与 EMD 和 EEMD 相比 CEEMDAN 分解后的分量变少，且各分量区别更大，分解结果更优。最近两年大多数时序预测的研究也都是利用了 CEEMDAN 来对时序进行分解处理。

三、LSTM 和 GRU 网络的内部结构和工作原理

（一）LSTM 神经网络

循环神经网络（RNN）已被广泛应用于语音识别等多领域中序列数据的建模过程。图表 4 展示了 RNN 按时间展开的链式结构，在这个循环结构中，主体结构 A 通过读取输入层信息 x_t 以及模型的上一时刻状态 h_{t-1} 来更新自身状态 h_t ，因此 RNN 能从之前的输入信息中获取有效信息，具有记忆功能，从而被应用于时序预测领域。

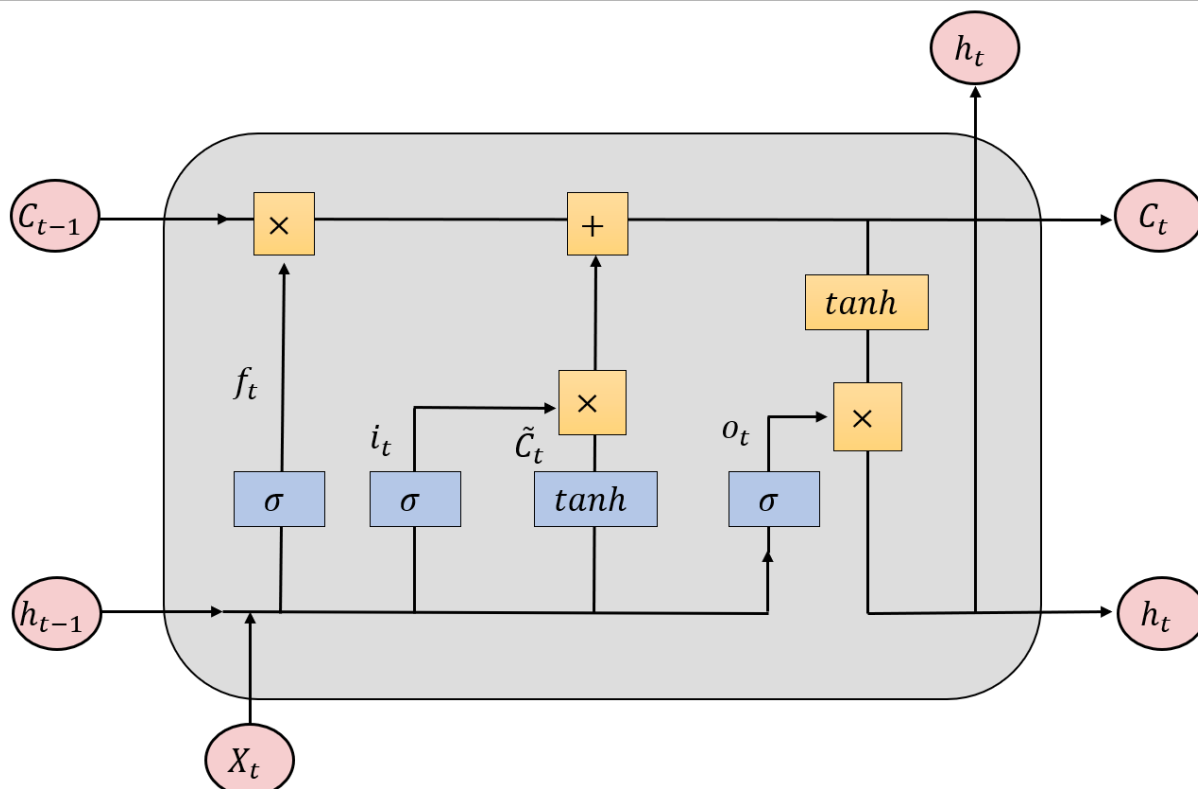
图表 4：单个 RNN 神经元结构和展开形式



资料来源：中信期货研究所

然而由于长期依赖问题，原始 RNN 隐藏单元的权重对网络的训练影响较大，会引起梯度消失和梯度爆炸现象。换一句话说，随着时间的推移，RNN 会忘记较长时间之前的信息，在一开始就遗漏重要信息，因此 RNN 只具有短时记忆。为了改善这个问题，LSTM 由 Hochreiter & Schmidhuber 在 1997 年被提出，随着 2012 年后深度学习的兴起，LSTM 的框架得到了进一步的优化，并在多个领域得到应用。

图表 5：LSTM 神经元结构

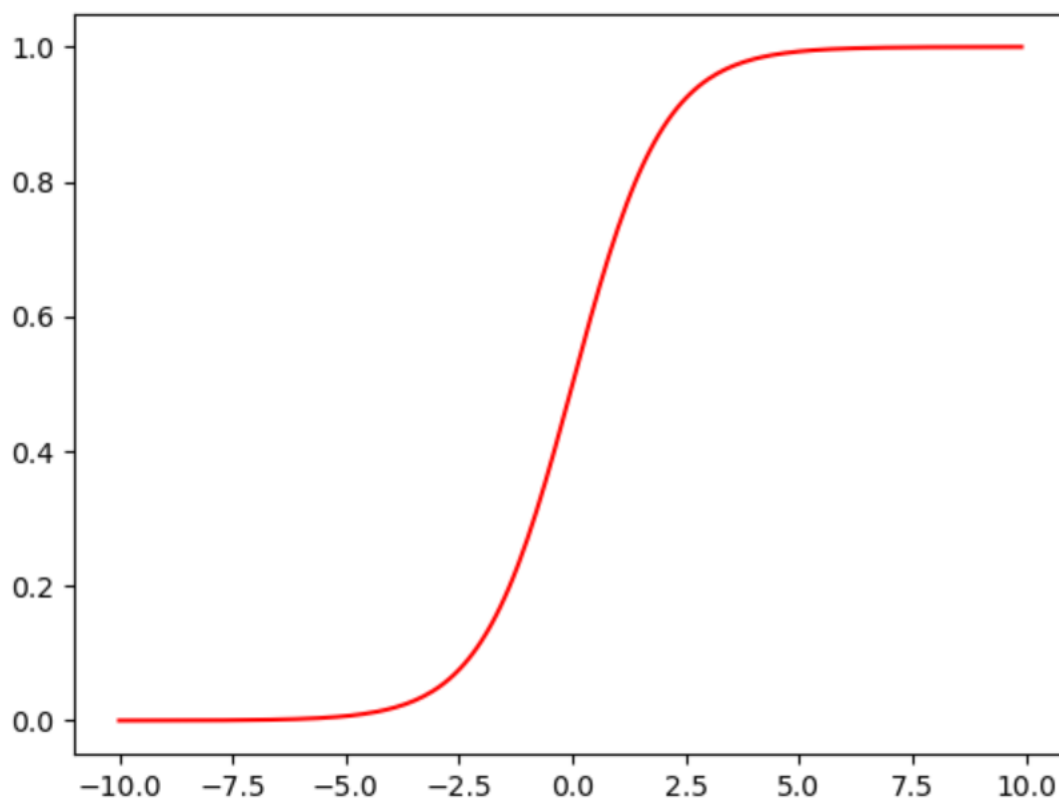


资料来源：中信期货研究所

图表 5 展示了 LSTM 神经网络记忆单元的基本结构，和传统 RNN 相比，LSTM 在保留了状态向量 h_t 的基础上，增加了能够保存长期状态的记忆状态向量 c_t （单元状态），同时引入了由遗忘门、输入门和输出门组成的门控机制来控制序列信息的遗忘与更新。我们先依次看一下每个“门”的逻辑和作用。首先是遗忘门，LSTM 通过遗忘门判断从当前状态中移除哪些信息，读取 h_{t-1} 和 c_t ，按公式 7 计算即可得出遗忘门向量 f_t ：

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (\text{公式 7})$$

图表 6：Sigmoid 函数图像



资料来源：中信期货研究所

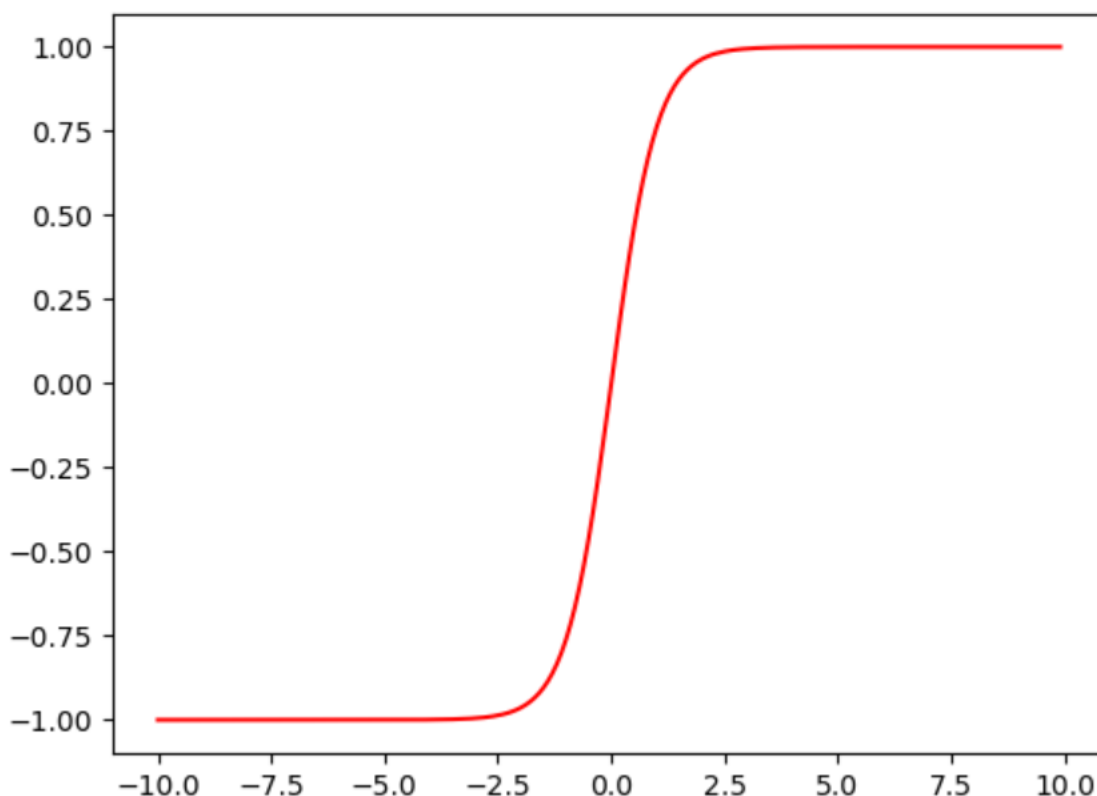
其中，激活函数 σ 表示 sigmoid 函数 ($S(x) = \frac{1}{1+e^{-x}}$)。图表 6 展示了 sigmoid 函数的图像，sigmoid 函数是生物学中常见的 S 型函数，它的值域范围为 $[0, 1]$ ，与概率值的范围相对应，经常被运用到神经网络中作为激活函数。因

此遗忘门向量 f_t 的每个元素都位于 $[0, 1]$ 范围内，0 表示完全不通过，1 表示完全通过。然后是输入门，输入门被用来判断哪些新信息会被更新到单元状态中，由一个 sigmoid 层和一个 tanh 层组成。输入门中的 sigmoid 层决定了需要更新的值，如公式 8 所示，通过线性变换和激活函数 σ 得到向量 i_t ；而 tanh 层则是为了创造出新的单元状态 \tilde{C}_t ，具体计算过程见公式 9。

$$i_t = \sigma (W_i[h_{t-1}, x_t] + b_i) \quad (\text{公式 8})$$

$$\tilde{C}_t = \tanh (W_c[h_{t-1}, x_t] + b_c) \quad (\text{公式 9})$$

图表 7：Tanh 函数图像



资料来源：中信期货研究所

其中，公式 9 中的 tanh 函数 ($\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$) 也是一种激活函数，图表 7 展示了它的图像，它的取值范围为 $[-1,1]$ ，有类似对数据去中心化的效果。sigmoid 函数取值范围为 $[0,1]$ 更适用于二分类任务，而 tanh 函数取值范围更加多样，在深度网络训练中模型的收敛速度会更优，因此在此处使用 tanh 函数。通过遗忘门和输入门，我们已知需要对信息做出怎么样的处理，接下来一步就是实现这些处理。如公式 10 所示，同时对应了图表 5 中最上端的那条水平线，将

旧单元状态 C_{t-1} 和遗忘门向量 f_t 相乘来丢弃需要丢弃的信息，再加上 $i_t \tilde{C}_t$ 以此向单元状态中添加需要更新的信息，最终得到新的单元状态 C_t 。

$$C_t = f_t C_{t-1} + i_t \tilde{C}_t \text{ (公式 10)}$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \text{ (公式 11)}$$

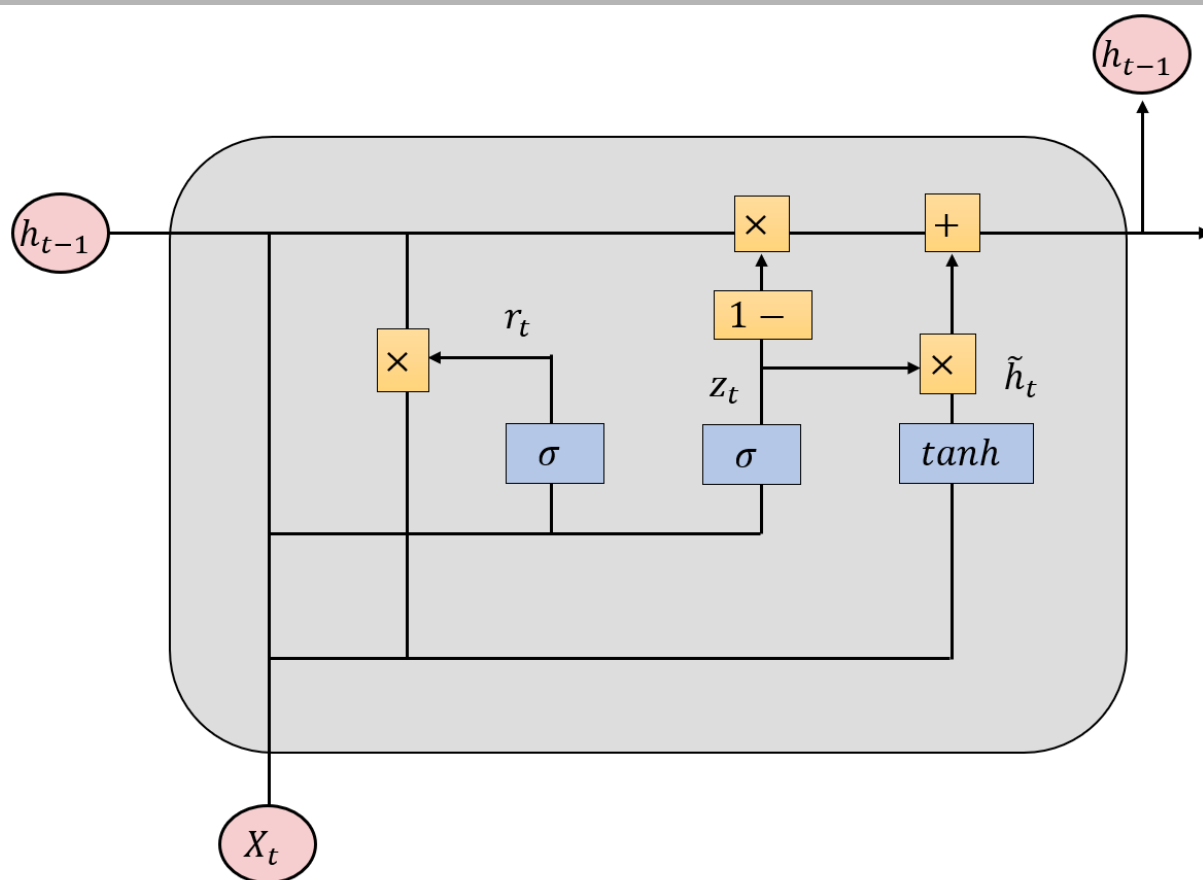
$$h_t = o_t \tanh(C_t) \text{ (公式 12)}$$

最后就是输出门，通过输出门来确定输出什么信息。如公式 11 和 12 所示，在输出门中，先通过 sigmoid 层确定输出什么信息，再使用 tanh 层对刚得出的新单元状态进行处理，最后将两者相乘得出最终输出信息 h_t 。复杂的神经元结构使得 LSTM 网络可以有效处理复杂的长期时序动态依赖关系，但也是因为复杂的结构使得 LSTM 在具体使用时过于繁琐，近年来众多学者提出了多种基于 LSTM 网络的变体，其中 GRU 网络就是最成功的变体之一，研究表明 GRU 结构更加简洁且保持和 LSTM 相似的预测效果。

（二）GRU 神经网络

GRU 神经网络由 Cho 等人在 2014 年提出。和 LSTM 相比，GRU 将三个门：遗忘门、输入门和输出门转变为两个门更新门 z_t 和重置门 r_t ，同时将 LSTM 的单元状态 C_t 和输出合并为一个状态 h_t 。

图表 8：GRU 神经元结构



资料来源：中信期货研究所

如图表 8 所示，GRU 网络只有两个输入值：此刻输入信息 X_t 和上一刻 GRU 输出值 h_{t-1} 和一个输出值：此刻 GRU 输出值 h_t 。现在，我们看一下 GRU 网络中两个门的作用。整体上理解公式 13 到 16 可以发现，更新门 z_t 是原本 LSTM 网络中遗忘门和输入门的结合，决定了 h_{t-1} 和 x_t 中有多少信息需要传递到当前隐藏状态 h_t 中。而重置门 r_t 决定了上一时刻隐藏状态 h_{t-1} 中有多少信息流入此刻的候选隐藏状态 \tilde{h}_t 中。如公式 16 所示，GRU 网络通过 $(1 - z_t) * h_{t-1}$ 对上一刻的隐藏状态选择性遗忘，同时通过 $z_t * \tilde{h}_t$ 对此刻的候选隐藏状态选择性记忆。

$$z_t = \sigma(W_z[h_{t-1}, x_t]) \text{ (公式 13)}$$

$$r_t = \sigma(W_r[h_{t-1}, x_t]) \text{ (公式 14)}$$

$$\tilde{h}_t = \tanh(W_{\tilde{h}}[r_t * h_{t-1}, x_t]) \text{ (公式 15)}$$

$$\hat{h}_t = (1 - Z_t) * \hat{h}_{t-1} + Z_t * \tilde{h}_t \text{ (公式 16)}$$

简化的门控结构使 GRU 网络的参数更少，训练速度加快，同时在很多情况中预测效果更优于 LSTM 网络。

四、组合模型构建

在上文中，各种分解算法和相关深度学习网络都得到了详细介绍，为了探究分解算法和深度学习的融合效果，本文选取 EMD 和 CEEMDAN 算法与 LSTM 网络和 GRU 网络各自组合形成组合模型并进行测试。

（一）IMF 重组

在分解算法的介绍中，我们有提到原始序列经分解算法分解后会得到相应不同频率的 IMF 分量和一个趋势项。为了降低组合模型的繁琐程度，我们将各 IMF 分量求均值进行 t 检验并重组，从而将 IMF 分量分为高低频两个子序列。结合最初的趋势项，原始序列就可以被重组成高频分量、低频分量和趋势项，具体步骤如下：

1. 将 IMF1 记为指标 1，IMF1+IMF2 为指标 2，以此类推，分别计算各指标的均值和标准差；
2. 给定显著性水平为 0.05，对刚刚得出的均值进行显著区别于 0 的 t 检验。其中，t 检验统计量为：

$$t = (\overline{IMFi} - 0) / \frac{\sigma_i}{\sqrt{n-1}} \text{ (公式 17)}$$

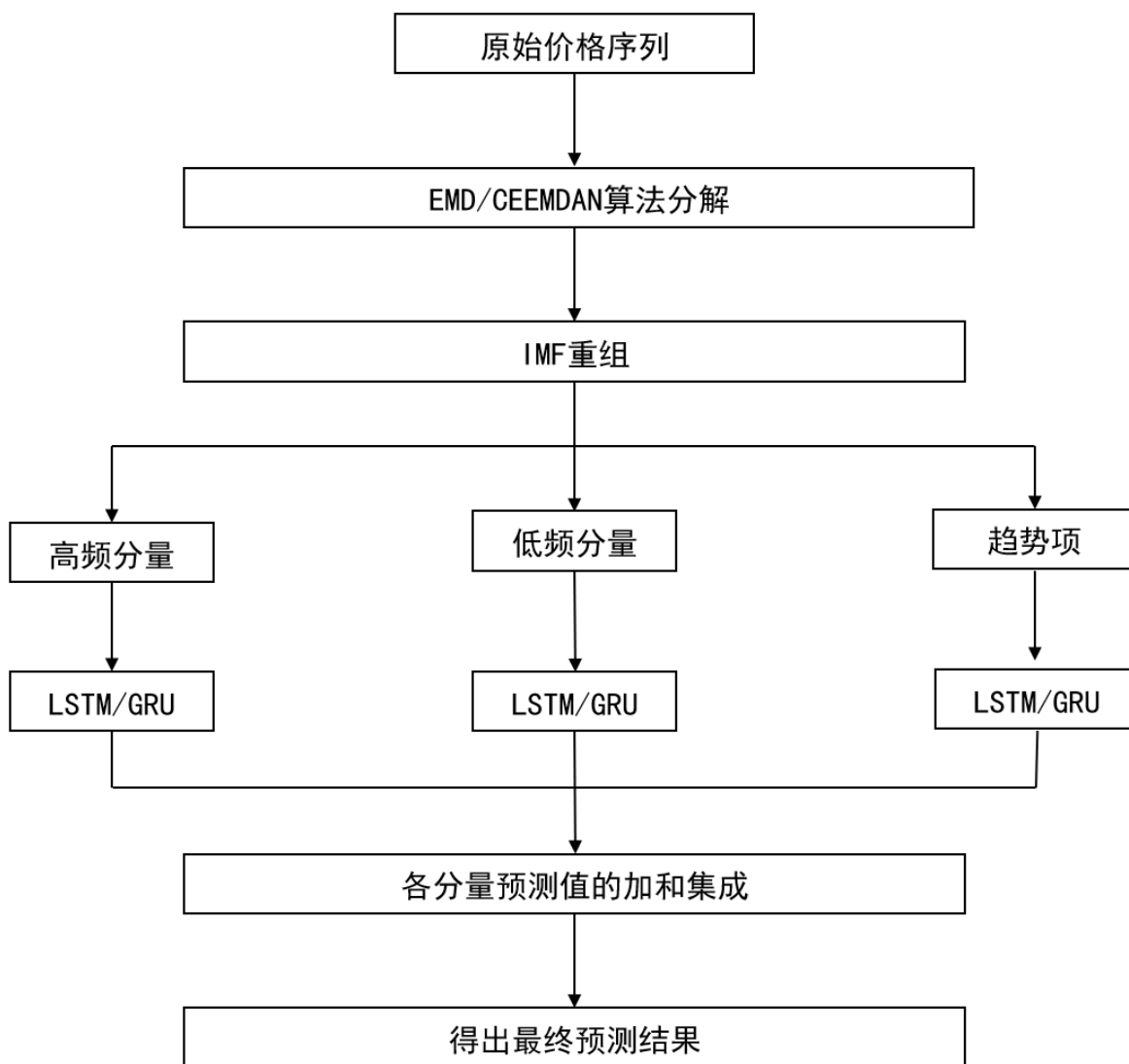
3. 若指标 i 为第一个均值显著不等于 0 的 IMF，则将 IMF1 至 IMF_{i-1} 加和得到原始序列的高频子序列，将 IMF_i 至 IMF_n 加和得到原始序列的低频子序列，并将分解剩余的 $r_n(t)$ 作为原始序列的趋势项。

（二）组合模型建模流程

图表 9 为组合模型的建模流程，模型主要包含三个部分，第一部分是通过 EMD 和 CEEMDAN 算法，将原始序列分解为多个 IMF 序列和残差序列，第二个部分是按照上文所述将各分量进行高低频和趋势项进行重组，第三部分是运用 LSTM

和 GRU 网络将重组后的子序列作为模型的输入数据进行训练预测，具体步骤如下：

图表 9：组合模型的流程



资料来源：中信期货研究所

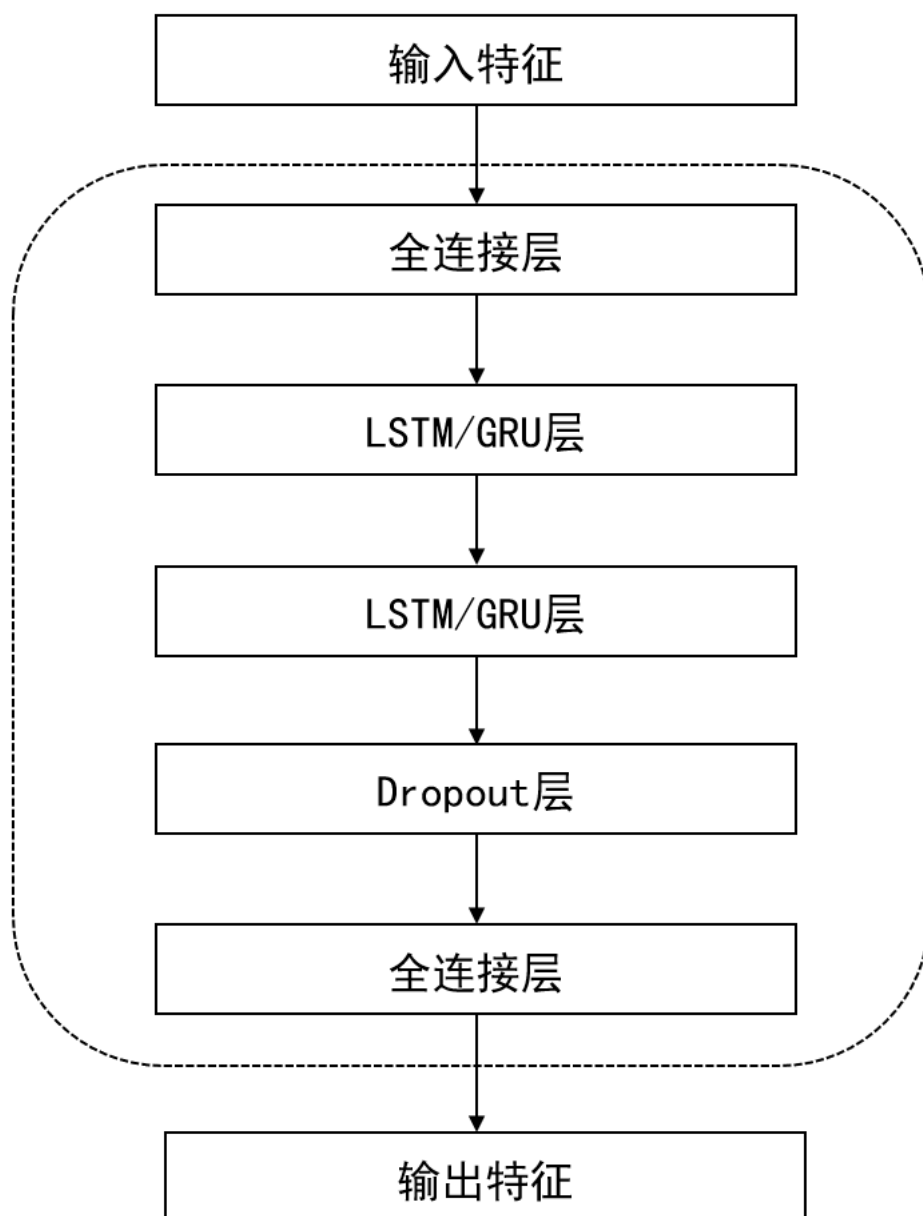
1. 运用 EMD 和 CEEMDAN 将原始序列 $s(t)$ 分解为多个 IMF 序列 $(IMF_i(t))$ 和残差 $r_N(t)$ 。
2. 按上文所述根据 t 检验将原始序列 $s(t)$ 重组为高频分量、低频分量和趋势项。
3. 针对重组后的高频分量、低频分量和趋势项分别进行 LSTM 和 GRU 训练预测。

4. 利用深度学习预测得出各序列的预测值，然后将预测值相加得到原始序列的最终预测值 $\tilde{S}(t)$ 。

（三）LSTM 和 GRU 网络的计算图结构和训练方法选择

图表 10 展示了本文构建的深层 LSTM/GRU 神经网络计算图结构，虚线方框内为神经网络结构。本文构建的深层网络的主体结构包括两层 LSTM/GRU 神经层和两层全连接层，参考过去文献将两个隐藏层的神经元个数设置为（8，8），Batch-size 设置为 64，模型迭代次数设置为 100。为了缓解过拟合问题，提高模型的泛化能力，在每层 LSTM/GRU 神经层后加入 Dropout 层，其中失活率设置为 0.2。本文采用 Mini-Batch 方法训练 LSTM 网络，选取均方误差（MSE）作为损失函数。为了提高训练收敛速率，采用 Adam 作为优化器进行优化训练。同时，使用早停机制，当连续 5 次迭代轮次的损失值出现增加时停止训练。本文采用的是循环预测建模方式，步长设置为 30，即以最近 30 个交易日的数据预测下一个交易日的收盘价。最后，本文是基于 Python 语言环境，以 PyTorch 作为深度学习框架进行训练和预测。

图表 10：LSTM/GRU 网络的计算图结构



资料来源：中信期货研究所

五、预测结果分析

（一）数据选择和处理

本文选取沪铜期货主力合约为主要测试标的，同时为了充分测试模型单步预测的泛化能力，选取沪深 300 股指期货和十年期国债期货主力合约为代表进行测试。考虑到上市时间，沪铜期货和沪深 300 股指期货数据集选用的时间范

围均为 2011 年 1 月 1 日至 2022 年 12 月 31 日，每个合约共 2917 行数据，十年期国债期货数据集选用时间为 2015 年 3 月 20 日至 2022 年 12 月 31 日，共 1898 行数据，将数据集划分为训练集、验证集和测试集，其中切割比率为 0.8:0.1:0.1。所有测试均使用日频数据，选取各主力合约的收盘价作为模型的输入特征。时序分解后不同序列大小差别较大，为缓解不同量纲带来的影响，提高模型训练效率，对数据进行 Z-score 标准化处理，转换公式如下：

$$x_{normalization} = \frac{x - \mu}{\sigma} \quad (\text{公式 18})$$

其中， μ 为原始数据的均值， σ 为原始数据的标准差。将归一化后的数据输入模型，模型转变为对归一化后的收盘价的预测，所以需要预测值进行去归一化处理来得到最终的预测结果。

（二）预测结果评价指标

本文选用 R-平方 (R^2)、均方根误差 (RMSE)、平均绝对误差 (MAE) 和方向正确率 (DA) 四种误差评价指标对模型预测性进行评估，四种指标的计算公式如下：

$$R^2 = 1 - \frac{\sum_{i=1}^m (y_i - \hat{y}_i)^2}{\sum_{i=1}^m (y_i - \bar{y})^2} \quad (\text{公式 19})$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (\text{公式 20})$$

$$MAE = \frac{1}{m} \sum_{i=1}^m |y_i - \hat{y}_i| \quad (\text{公式 21})$$

$$DA = \left(1 - \frac{\sum_{i=1}^{n-1} |\text{sign}(y'_{i+1} - y_i) - \text{sign}(y_{i+1} - y_i)|}{n-1} \right) \times 100\% \quad (\text{公式 22})$$

其中， y_i 为真实值， \hat{y}_i 为预测值， \bar{y} 为真实值的平均值， m 为序列长度。 R^2 用来衡量模型的拟合优度，越接近 1 表明模型拟合程度越好。RMSE 和 MAE 用来衡量真实值和预测值间的偏差，值越小表明预测误差越小。DA 衡量了模型在序列方向上的准确性，值越高代表方向准确率越高。

（三）预测结果展示和分析

首先，我们将 CU 的收盘价通过 EMD 和 CEEMDAN 进行分解并重组，形成各自的高低频和趋势序列，利用已建立的 LSTM 模型对各序列进行预测。经 EMD 和 CEEMDAN 分解后发现，低频序列中 IMF 分量较少甚至没有，因此将低频序列和趋势项合并考虑。如图表 11 所示，经模型预测后发现，模型对纯趋势项的学习能力较弱，因此需测试后对 IMF 分量进行重新重组。同时高频序列中含有大量无序噪音，预测精度难以提升，因此尝试通过剔除噪音含量过多的 IMF 分量来对高频序列进行优化。

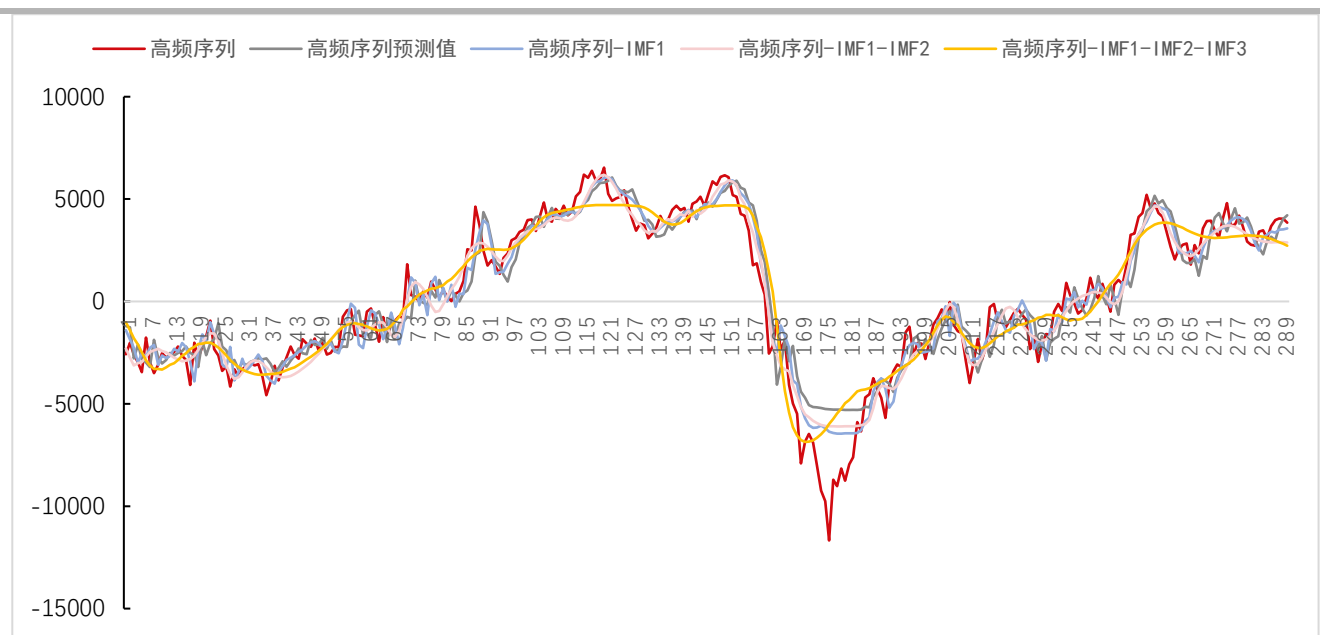
图表 11：CU 分解后各频率序列 LSTM 模型预测效果比较

序列名称	R^2	RMSE	MAE	DA
EMD_高频序列	88.66%	1231.09	903.12	49.13%
EMD_低频序列	99.64%	305.17	285.87	3.81%
EMD_趋势项	-165.48%	2063.68	1791.35	0.34%
EMD_低频序列+趋势项	98.63%	454.57	405.14	20.07%
CEEMDAN_高频序列	91.33%	1079.57	850.94	50.17%
CEEMDAN_低频序列	99.19%	455.28	430.52	5.19%
CEEMDAN_趋势项	-9072.86%	11546.41	11482.64	0.35%
CEEMDAN_低频序列+趋势序列	16.87%	3532.21	2376.00	58.48%

资料来源：同花顺 iFind、中信期货研究所

如图表 11 所示，经 EMD 分解后的低频序列和趋势项的组合序列预测模型的 R^2 已经达到 98.63%，预测效果已经比较优异，因此不对此组合序列进行重新重组，直接对高频序列进行去噪优化。

图表 12：CU 经 EMD 分解高频序列模型预测结果展示



资料来源：同花顺 iFind、中信期货研究所

图表 13：CU 经 EMD 分解后高频率序列模型预测效果比较

序列名称	R^2	RMSE	MAE	DA
高频率序列	88.66%	1231.09	903.12	49.13%
高频率序列-IMF1	93.69%	918.09	677.14	60.90%
高势序列-IMF1-IMF2	93.92%	901.17	651.89	61.94%
高势序列-IMF1-IMF2-IMF3	91.01%	1096.39	801.53	57.79%

资料来源：同花顺 iFind、中信期货研究所

图表 12 和 13 展示了通过剔除相应高频 IMF 分量后高频率序列模型的预测结果，可以发现剔除 IMF1 和剔除 IMF1+IMF2 可以显著提高模型的拟合效果， R^2 提升了 5%以上，RMSE 和 MAE 也都缩小了 25%以上。剔除 IMF1+IMF2 的优化效果稍许优于剔除 IMF1，因此选择剔除 IMF1 和 IMF2 两个分量来形成高频率序列模型的新的输入特征。图表 11 也展示了 CU 经 CEEMDAN 分解后各序列模型的预测效果，可以发现趋势项的预测效果很差，因此需要对 IMF 分量进行重新重组。

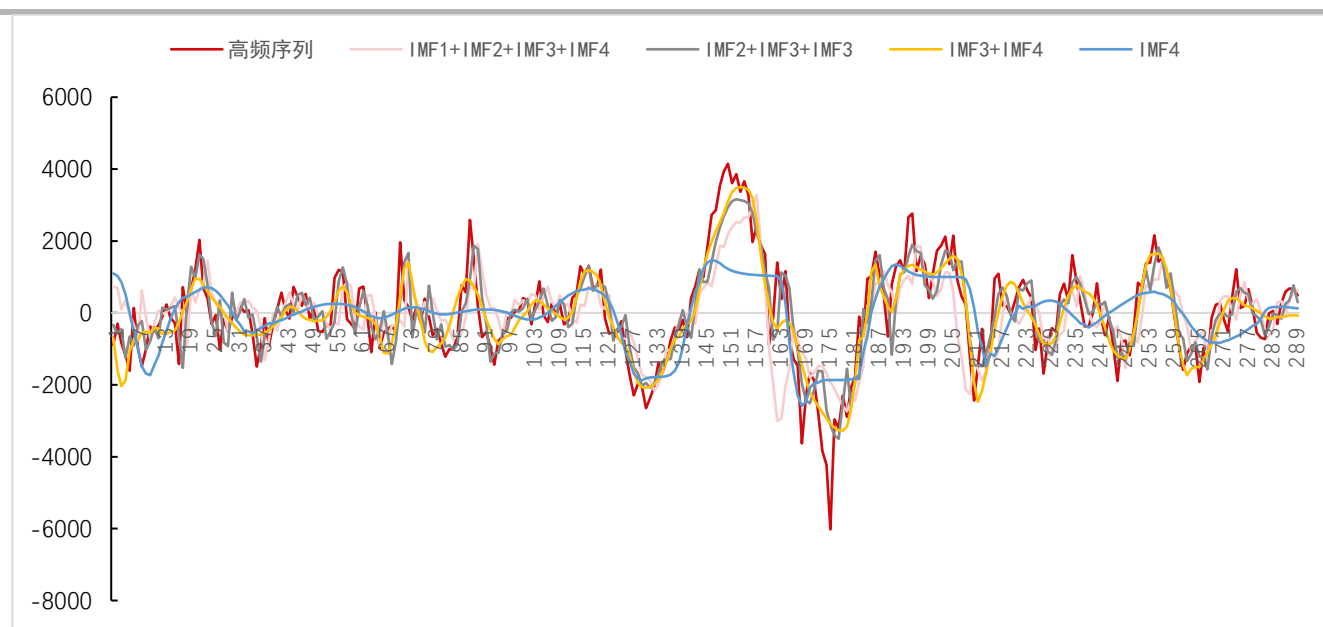
图表 14：CU 经 CEEMDAN 分解各序列 LSTM 模型预测效果比较

序列名称	R^2	RMSE	MAE	DA
低频+趋势序列（趋势+IMF7）	12.77%	3543.18	2595.24	61.59%
低频+趋势序列（趋势+IMF7+IMF6）	90.87%	1409.06	1146.20	23.53%
低频+趋势序列（趋势+IMF7+IMF6+IMF5）	95.68%	969.28	784.28	28.03%
低频+趋势序列（趋势+IMF7+IMF6+IMF5+IMF4）	93.70%	1222.08	847.18	33.56%
高势序列（IMF1+IMF2+IMF3+IMF4）	44.91%	1004.68	750.21	57.09%
高势序列（IMF2+IMF3+IMF4）	75.07%	675.81	514.93	64.36%
高势序列（IMF3+IMF4）	77.91%	636.17	494.91	67.82%
高势序列（IMF4）	50.24%	964.00	727.39	60.90%

资料来源：同花顺 iFind、中信期货研究所

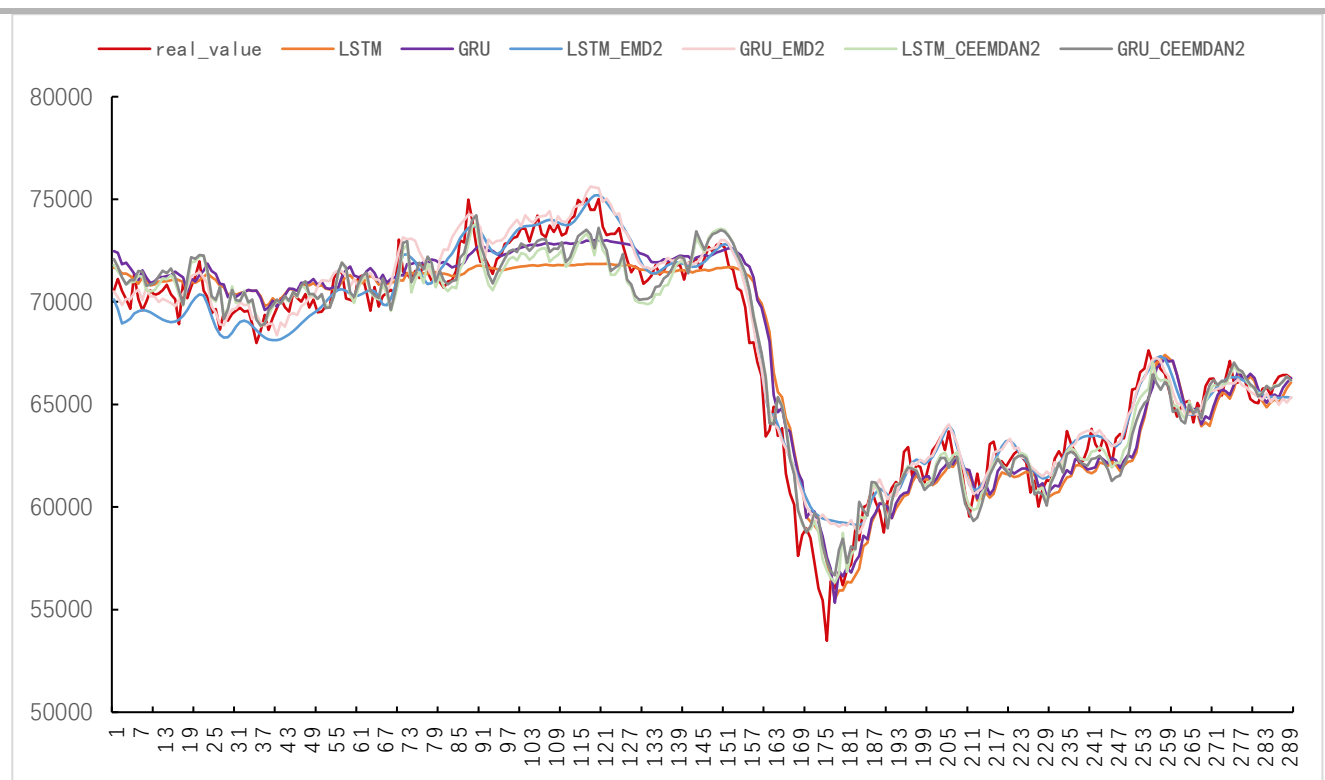
如图表 14 显示，通过将趋势项和各低频率 IMF 分量分别组合进行预测后，发现趋势项、IMF7、IMF6 和 IMF5 组合后预测值的拟合效果最佳，因此挑选此组合形式形成全新组合序列，同时将剩余高频度 IMF 分量相加形成高频率序列。和 EMD 一样，需对高频率序列进行去噪操作，通过不断剔除高频度 IMF 分量进行预测检验。如图表 14 和图表 15 所示，IMF3 和 IMF4 组合后对原高频率序列预测效果最佳，和不去噪预测相比， R^2 提高了 33%，RMSE 和 MAE 分别缩小了 37% 和 34% 左右。最后，我们对重组和去噪后形成的全新高频率序列和组合序列进行不同深度学习模型的预测。

图表 15：CU 经 CEEMDAN 分解高频率序列模型预测结果展示



资料来源：同花顺 iFind、中信期货研究所

图表 16：CU 各类模型预测结果展示



资料来源：同花顺 iFind、中信期货研究所

图表 17：CU 各类模型预测效果比较

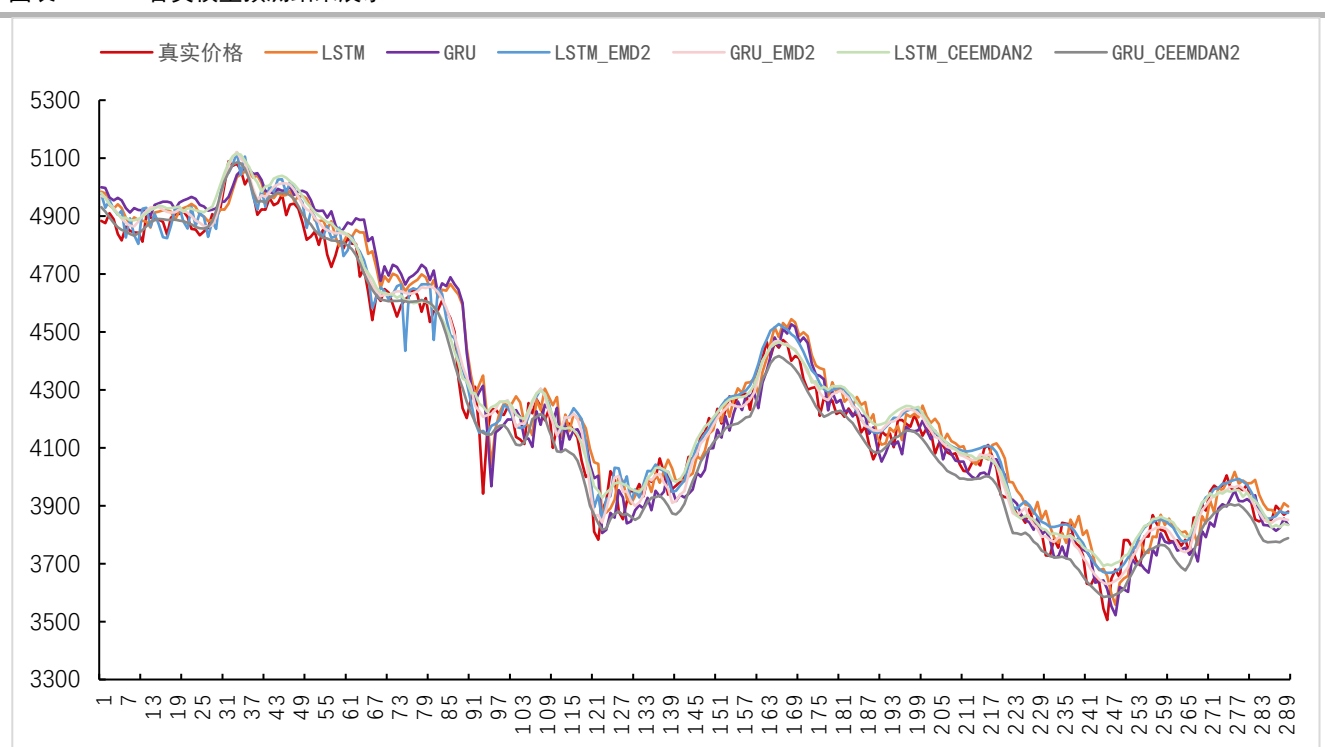
模型名称	R^2	RMSE	MAE	DA
LSTM	91.01%	1467.27	1142.69	51.90%
GRU	92.52%	1339.08	1032.62	52.25%
EMD_LSTM1	92.41%	1348.70	957.24	50.87%
EMD_LSTM2	95.31%	1059.43	779.43	55.71%
EMD_GRU1	93.00%	1294.65	949.29	52.60%
EMD_GRU2	95.54%	1033.25	747.80	63.32%
CEEMDAN_LSTM1	94.07%	1191.98	912.06	51.56%
CEEMDAN_LSTM2	95.71%	1013.82	805.46	59.17%
CEEMDAN_GRU1	94.06%	1193.39	937.84	54.33%
CEEMDAN_GRU2	95.41%	1049.14	838.76	54.67%

资料来源：同花顺 iFind、中信期货研究所

图表 16 和 17 展示了 CU 在各模型下的预测结果，图表中模型名称的 1 和 2 分别带别了纯重组模型和重组加去噪模型。可以发现不管是经 EMD 还是 CEEMDAN 分解，分解后的模型预测效果都比初始数据预测效果提高很多。同时，去噪后的模型预测效果也要比纯重组的模型优越。综合来看，重组加去噪模型的 R^2 比纯深度学习模型提高了 4% 左右。各组和中 CEEMDAN_LSTM2 模型预测的拟合效果最优， R^2 达到 95.71%，方向正确率达到 59.17%。虽然 EMD_GRU2 模型的拟合程度略低于 CEEMDAN_LSTM2 模型，但是它的方向正确率是最高的，达到 63.32%。

为了检测模型的泛化能力，将同样的重组去噪深度学习模型运用在 IF 期货和 T 期货上，图表 18~21 展示了各自的预测结果。和 CU 的预测结果一样，IF 和 T 的预测结果也证明了重组去噪模型可以较大提高原本深度学习模型的拟合程度和方向正确率。IF 预测效果较好的模型有 EMD_GRU2 和 CEEMDAN_GRU2， R^2 分别达到 98.43% 和 97.54%，方向正确率分别达到 63.32% 和 64.36%。T 预测效果较好的模型和 IF 一样同为 EMD_GRU2 和 CEEMDAN_GRU2， R^2 分别达到 92.52% 和 92.63%，方向正确率分别达到 65.78% 和 66.31%。因此，我们认为分解算法可以提升 LSTM 和 GRU 神经网络的学习能力，降低噪音对模型的干扰，降低模型预测的滞后性。

图表 18：IF 各类模型预测结果展示



资料来源：同花顺 iFind、中信期货研究所

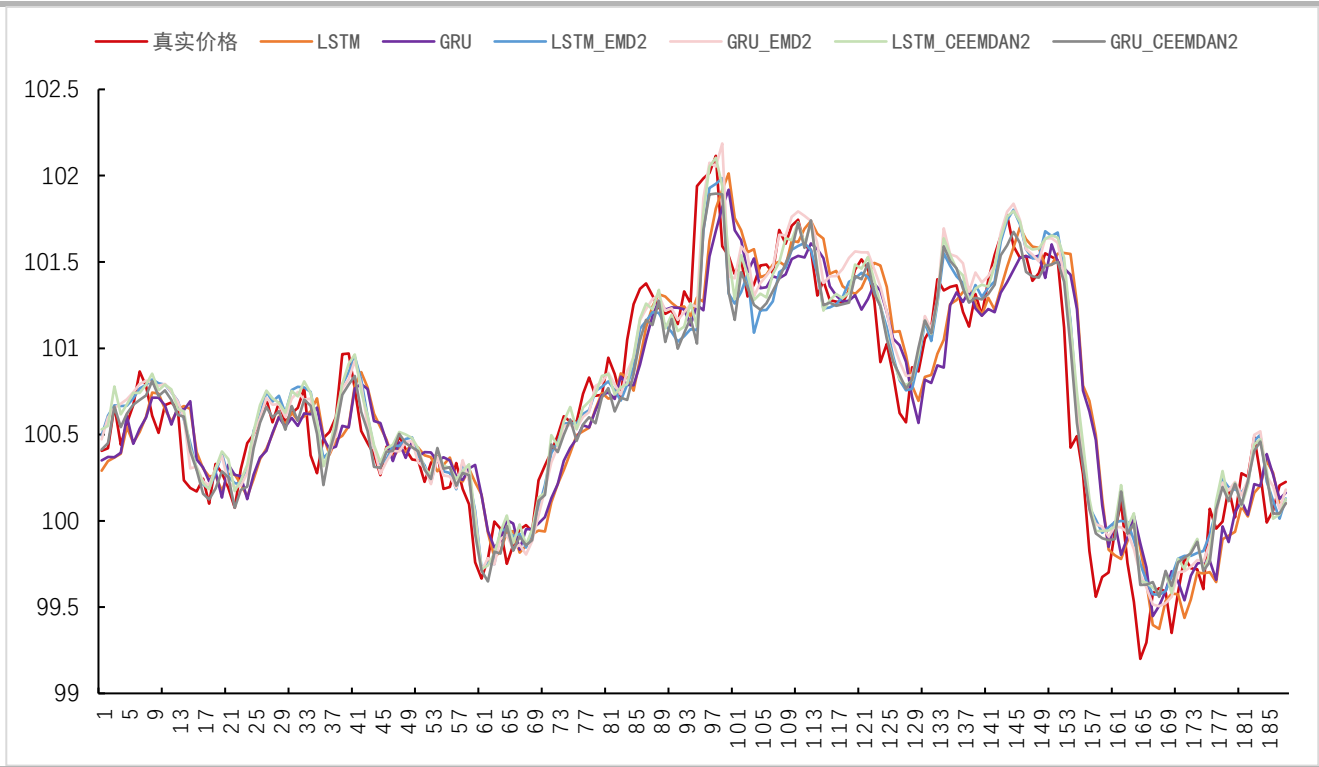
图表 19：IF 各类模型预测效果比较

模型名称	R^2	RMSE	MAE	DA
LSTM	95.35%	86.92	66.12	48.79%
GRU	94.78%	92.10	71.86	51.90%
EMD_LSTM1	96.83%	71.76	52.49	54.67%
EMD_LSTM2	97.93%	57.93	45.51	60.90%
EMD_GRU1	96.21%	78.38	62.41	58.82%
EMD_GRU2	98.43%	50.39	38.48	63.32%
CEEMDAN_LSTM1	95.85%	82.15	64.42	51.56%
CEEMDAN_LSTM2	97.76%	60.31	46.89	58.48%
CEEMDAN_GRU1	96.09%	79.70	63.77	59.86%

CEEMDAN_GRU2	97.54%	63.26	52.14	64.36%
--------------	--------	-------	-------	--------

资料来源：同花顺 iFind、中信期货研究所

图表 20：T 模型各类预测结果展示



资料来源：同花顺 iFind、中信期货研究所

图表 21：T 模型各类预测效果比较

模型名称	R^2	RMSE	MAE	DA
LSTM	81.34%	0.2781	0.2125	48.13%
GRU	82.28%	0.2710	0.2036	51.34%
EMD_LSTM1	85.07%	0.2487	0.1885	54.01%
EMD_LSTM2	91.92%	0.1830	0.1384	64.71%
EMD_GRU1	84.31%	0.2550	0.1910	55.08%
EMD_GRU2	92.52%	0.1761	0.1319	65.78%
CEEMDAN_LSTM1	85.60%	0.2443	0.1854	54.01%
CEEMDAN_LSTM2	92.43%	0.1771	0.1323	64.84%
CEEMDAN_GRU1	83.28%	0.2632	0.2026	50.80%
CEEMDAN_GRU2	92.63%	0.1748	0.1314	66.31%

资料来源：同花顺 iFind、中信期货研究所

六、结论与展望

本文将分解算法（EMD 和 CEEMDAN）和深度学习算法（LSTM 和 GRU）分别组合对 CU、IF 和 T 三种期货品种的日度收盘价进行单步预测后，发现：

1. 分解算法（EMD 和 CEEMDAN）将初始价格序列分解为多个 IMF 分量和趋势项后，不能简单依靠 T 检验进行重组形成高频、低频和趋势项。LSTM 和 GRU 对纯趋势项学习能力较弱，若想达到最优效果，需先将 IMF 分量和趋势项分别组合测试进行挑选，再将剩余的 IMF 分量加成为高频率序列。
2. 初始的高频率 IMF 分量含有大量的噪音，会对神经网络预测进行干扰，因此可以通过剔除部分高频率 IMF 分量来对分解后的高频率序列进行预测。
3. 分解算法（EMD 和 CEEMDAN）均能对深度学习算法（LSTM 和 GRU）从拟合效果和方向准确性上起到相应的优化效果。去噪后的模型预测效果最优，在三个品种的预测中，方向准确率均可达到 60%以上。
4. GRU 的预测效果不是一定优于 LSTM 的预测效果，一般情况下，两种算法的预测效果是相似的，GRU 略优于 LSTM。
5. EMD 和 CEEMDAN 对 GRU 和 LSTM 的优化程度是不确定的，虽然都会对原始算法起到优化效果，但不代表 CEEMDAN 对深度学习算法的优化强度一定高于 EMD。

考虑到本篇为此系列的第一篇，只介绍了 RNN 系列的深度学习算法，在之后的报告中会对其他系列的算法进行详细介绍。同时，本次测试并未对神经网络层数、神经元个数和迭代次数进行过多讨论，后续可以通过更改参数来对模型进行更深地挖掘。

参考文献

- 1) M T Rosenstein, J J Collins, C J De Luca. A Practical Method for Calculating Largest Lyapunov Exponents from Small Data Sets III[J]. Physica D, 1993, (65).
- 2) 高红兵, 潘瑾, 陈宏民. 我国证券市场混沌的判断[J]. 系统工程, 2000, 18(6).
- 3) 刘海飞, 李心丹. 基于 EMD 方法的股票价格预测与实证研究[J]. 统计与决策, 2010(23).
- 4) 齐绍洲, 赵鑫, 谭秀杰. 基于 EEMD 模型的中国碳市场价格形成机制研究[J]. 2015, 68(4).
- 5) 郝欢, 王华力, 魏勤. 经验模态分解理论及其应用[J]. 高技术通讯. 2016, 26(1):67-80.
- 6) Hochreiter S, Schmidhuber J. Long Short-term Memory[J]. Neural Computation, 1997(8):1735-1780.
- 7) Thomas F, Christopher K. Deep Learning with Long short-term Memory Networks for Financial Market Predictions[J]. European Journal of Operational Research, 2018, 270(2):654-669.
- 8) 杨青, 王晨蔚. 基于深度学习 LSTM 神经网络的全球股票指数预测研究[J]. 统计研究. 2019, 36(3).
- 9) 姚洪刚, 沐年国. EMD-LSTM 模型对金融时间序列的预测[J]. 计算机工程与应用. 2021, 57(5).
- 10) 张倩玉, 严冬梅, 韩佳彤. 结合深度学习和分解算法的股票价格预测研究[J]. 计算机工程与应用. 2021, 57(5).
- 11) 谷丽琼, 吴运杰, 逢金辉. 基于 Attention 机制的 GRU 股票预测模型[J]. 系统工程. 2020, 58(12).
- 12) 贺毅岳, 李萍, 韩进博. 基于 CEEMDAN-LSTM 的股票市场指数预测建模研究[J]. 统计与信息论坛. 2020, 35(6).

免责声明

除非另有说明，中信期货有限公司拥有本报告的版权和/或其他相关知识产权。未经中信期货有限公司事先书面许可，任何单位或个人不得以任何方式复制、转载、引用、刊登、发表、发行、修改、翻译此报告的全部或部分材料、内容。除非另有说明，本报告中使用的所有商标、服务标记及标记均为中信期货有限公司所有或经合法授权被许可使用的商标、服务标记及标记。未经中信期货有限公司或商标所有权人的书面许可，任何单位或个人不得使用该商标、服务标记及标记。

如果在任何国家或地区管辖范围内，本报告内容或其适用与任何政府机构、监管机构、自律组织或者清算机构的法律、规则或规定内容相抵触，或者中信期货有限公司未被授权在当地提供这种信息或服务，那么本报告的内容并不意图提供给这些地区的个人或组织，任何个人或组织也不得在当地查看或使用本报告。本报告所载的内容并非适用于所有国家或地区或者适用于所有人。

此报告所载的全部内容仅作参考之用。此报告的内容不构成对任何人的投资建议，且中信期货有限公司不会因接收人收到此报告而视其为客户。

尽管本报告中所包含的信息是我们于发布之时从我们认为可靠的渠道获得，但中信期货有限公司对于本报告所载的信息、观点以及数据的准确性、可靠性、时效性以及完整性不作任何明确或隐含的保证。因此任何人不得对本报告所载的信息、观点以及数据的准确性、可靠性、时效性及完整性产生任何依赖，且中信期货有限公司不对因使用此报告及所载材料而造成的损失承担任何责任。本报告不应取代个人的独立判断。本报告仅反映编写人的不同设想、见解及分析方法。本报告所载的观点并不代表中信期货有限公司或任何其附属或联营公司的立场。

此报告中所指的投资及服务可能不适合阁下。我们建议阁下如有任何疑问应咨询独立投资顾问。此报告不构成任何投资、法律、会计或税务建议，且不担保任何投资及策略适合阁下。此报告并不构成中信期货有限公司给予阁下的任何私人咨询建议。

深圳总部

地址：深圳市福田区中心三路8号卓越时代广场（二期）北座13层1301-1305、14层

邮编：518048

电话：400-990-8826

传真：(0755) 83241191

网址：<http://www.citicsf.com>