

2025年08月28日

证券研究报告·金融工程专题报告

机器学习应用系列



西南证券
SOUTHWEST SECURITIES

DAFAT：基于 Transformer 模型的自适应解决方案

摘要

- 传统 Transformer 模型局限性与 DATAT 模型概述：**Transformer 模型在位置编码以及注意力机制方面均有一定的优化空间，同时大多数深度学习模型均以高频量价为主，基本面特征频次低且更新滞后难以应用。经回测，在仅选择量价特征的基础上，Transformer 模型 2019 年 1 月至 2025 年 7 月 IC 均值 9.42%，多头组合年化收益率为 28.70%，月均单边换手率 0.88X，具备一定的提升空间。
- 动态位置编码自适应：**针对 Transformer 模型传统正余弦位置编码存在的周期性失配以及状态盲区等问题，本文提出了双通道门控动态位置编码（Double-Gate DPE），分别包括时间周期编码以及市场状态编码，代替原有正余弦位置编码。经回测 DPE_Transformer 因子 2019 年 1 月至 2025 年 7 月 IC 均值为 9.54%，多头组合年化收益率 30.60%，月均单边换手率 0.86X。
- 稀疏注意力机制自适应：**针对传统 Transformer 全连接中存在的计算复杂度高以及噪声敏感等问题，本文提出了提出三重稀疏化注意力自适应机制，该机制可以在提升计算效率的同时，加大注意力机制的局部针对性聚焦模式，减少对噪声时段的关注。该稀疏注意力机制分别包括市场波动率门控、局部注意力窗口以及 TopK 稀疏选择。经回测，SA_Transformer 模型 2019 年 1 月至 2025 年 7 月 IC 9.98%，多头组合年化收益率为 29.64%，月均单边换手率 0.87X。
- 多尺度时序信息融合自适应：**除量价特征以外，本文引入基本面特征以补全特征维度。但传统基本面特征季度发布，与日频量价特征存在显著频次差异。根据常规量价特征与基本面特征处理存在的问题，本文设计了多尺度信息融合框架（微观、中观以及宏观）。经回测，MF_Transformer 因子 2019 年 1 月至 2025 年 7 月 IC 10.09%，多头组合年化收益率 29.63%，月均单边换手率 0.82X。
- 模型消融测试：**上述三个不同的自适应优化模块（DPE、SA、MF）优化情况下模型较基础 Transformer 模型均有所提升，但两两组合的情况下，模型表现相较于传统 Transformer 模型具备一定的削弱效果。三个自适应模块组合下 DAFAT 自 2019 年 1 月至 2025 年 7 月因子 IC 均值 11.07%，多头组合年化收益率 32.30%，最大回撤率 33.00%，整体较其他模型表现提升相对明显。
- 指数增强策略：**DAFAT 模型沪深 300 指数增强策略年化超额收益率 12.65%，近一个月、三个月、一年、三年超额收益率分别为 -0.25%、0.20%、14.76% 以及 10.39%；DAFAT 模型中证 1000 指数增强策略年化超额收益率 14.57%，近一个月、三个月、一年、三年超额收益率分别为 -0.60%、2.41%、20.85% 以及 16.79%。
- 风险提示：**报告对应的相关结论完全基于公开的历史数据进行算法构建、统计以及计算，文中部分数据有一定滞后性，同时也存在第三方数据提供不准确或者缺失等风险；策略效果结论仅针对于回测区间得出，并不预示其未来表现，也不能保证未来的可持续性，亦不构成投资收益的保证或投资建议。

西南证券研究院

分析师：郑琳琳

执业证号：S1250522110001

邮箱：zhengll@swsc.com.cn

联系人：祝晨宇

邮箱：zhcy@swsc.com.cn

相关研究

- 加权影线频率与 K 线形态因子 (2025-08-28)
- 可转债 K 线看跌信号与交易增强策略——量化方法在债券研究中的应用四 (2025-06-09)
- 基于历史 K 线形态的因子选股研究 (2025-05-26)
- 科技成长产业变革趋势下基金产品投资策略评价与优选 (2025-05-26)
- PINN 信息约束与时序截面双流网络选股模型 (2025-04-09)
- 基于产业链和交易结构的豆粕期货择时框架 (2025-03-19)
- 可转债 K 线技术分析与 K 线形态因子——量化方法在债券研究中的应用三 (2025-03-14)
- 识时通变：宏微同频成长价值风格轮动策略 (2025-01-10)
- 超额收益如何回归？——2025 年金融工程策略报告 (2025-01-06)
- ESG 投资概述及其在多因子策略中的应用 (2024-12-11)

目 录

1 传统 Transformer 模型局限性及可优化点	1
1.1 Transformer 模型概述.....	1
1.2 Transformer 与传统序列模型对比.....	4
1.3 Transformer 主要优势.....	5
1.4 传统 Transformer 模型局限性及可优化点.....	6
2 位置编码自适应：动态位置编码解决方案	10
2.1 动态位置编码自适应 Transformer 模型回测结果.....	11
3 注意力自适应：稀疏注意力解决方案	14
3.1 稀疏注意力自适应 Transformer 模型回测结果.....	15
4 信息融合自适应：多尺度信息融合解决方案	18
4.1 信息融合自适应 Transformer 模型测试.....	19
5 模型消融测试	22
6 指数增强策略	25
6.1 沪深 300 指数增强策略.....	25
6.2 中证 1000 指数增强策略.....	27
7 总结与展望	30
8 风险提示	30

图 目 录

图 1: TransFormer 网络结构.....	1
图 2: 注意力机制	2
图 3: 多头注意力示意	3
图 4: 残差连接示意.....	3
图 5: 基础 Transformer 因子 IC.....	7
图 6: 基础 Transformer 因子分组-绝对净值	7
图 7: 基础 TransFormer 因子分组测试-相对净值.....	8
图 8: 基础 Transformer 因子多头分年度净值	9
图 9: 动态位置编码自适应: 时间周期编码+市场状态编码	10
图 10: DPE_Transformer 因子 IC.....	12
图 11: DPE_Transformer 因子分组-绝对净值.....	12
图 12: DPE_Transformer 因子分组测试-相对净值.....	13
图 13: DPE_Transformer 因子多头分年度净值.....	13
图 14: 注意力编码自适应: 波动率门控+局部注意力窗口+Top-k 稀疏选择.....	15
图 15: SA_Transformer 因子 IC.....	16
图 16: SA_Transformer 因子分组-绝对净值	16
图 17: SA_Transformer 因子分组测试-相对净值	17
图 18: SA_Transformer 因子多头分年度净值	17
图 19: 多尺度信息融合自适应: 微、中、宏观尺度做基本面及量价信息处理.....	18
图 20: MF_Transformer 因子 IC	20
图 21: MF_Transformer 因子分组-绝对净值	20
图 22: MF_Transformer 因子分组测试-相对净值	21
图 23: MF_Transformer 因子多头分年度净值.....	21
图 24: DAFAT 模型结构示意.....	22
图 25: Transformer 自适应模块消融测试-相对净值.....	23
图 26: DAFAT 因子 IC.....	23
图 27: DAFAT 因子分组-绝对净值.....	23
图 28: DAFAT 因子分组测试-相对净值.....	24
图 29: DAFAT 因子多头分年度净值.....	25
图 30: DAFAT 沪深 300 指数增强表现.....	26
图 31: DAFAT 沪深 300 指数增强超额滚动回撤	26
图 32: DAFAT 沪深 300 指数增强表现 (近一月)	27
图 33: DAFAT 沪深 300 指数增强表现 (近三月)	27
图 34: DAFAT 沪深 300 指数增强表现 (近一年)	27
图 35: DAFAT 沪深 300 指数增强表现 (近三年)	27
图 36: DAFAT 中证 1000 指数增强表现.....	28
图 37: DAFAT 中证 1000 指数增强相对净值.....	28
图 38: DAFAT 中证 1000 指数增强表现 (近一月)	29
图 39: DAFAT 中证 1000 指数增强表现 (近三月)	29
图 40: DAFAT 中证 1000 指数增强表现 (近一年)	29
图 41: DAFAT 中证 1000 指数增强表现 (近三年)	29

表 目 录

表 1: Transformer 模型各结构功能及创新.....	3
表 2: 不同模型计算复杂度性能特点对比.....	4
表 3: 不同模型梯度传播特性与长程依赖建模能力.....	4
表 4: Transformer 模型量价时序特征列表.....	6
表 5: 基础 Transformer 回测结果.....	7
表 6: 基础 Transformer 因子 10 分组表现.....	8
表 7: 基础 Transformer 因子多头分年度超额表现.....	9
表 8: DPE_Transformer 回测结果.....	11
表 9: DPE_Transformer 因子 10 分组表现.....	12
表 10: DPE_Transformer 因子多头分年度超额表现.....	13
表 11: SA_Transformer 因子回测结果.....	15
表 12: SA_Transformer 因子分组表现.....	16
表 13: SA_Transformer 因子多头分年度超额表现.....	17
表 14: 模型基本面特征列表.....	19
表 15: MF_Transformer 因子回测结果.....	20
表 16: MF_Transformer 因子分组表现.....	20
表 17: MF_Transformer 因子多头分年度超额表现.....	21
表 18: 自适应优化模块模型消融测试.....	22
表 19: DAFAT 因子分组表现.....	24
表 20: DAFAT 因子多头分年度超额表现.....	25
表 21: 沪深 300 指数成分股因子测试.....	25
表 22: 沪深 300 指数增强策略回测结果.....	26
表 23: DAFAT 沪深 300 指数增强策略分区间回测结果.....	26
表 24: 中证 1000 指数成分股因子测试.....	28
表 25: 中证 1000 指数增强策略回测结果.....	28
表 26: DAFAT 中证 1000 指数增强策略分区间回测结果.....	29

1 传统 Transformer 模型局限性及可优化点

1.1 Transformer 模型概述

Transformer 模型于 2017 年由 Google 的研究团队在论文《Attention Is All You Need》中首次提出。该论文的主要作者包括 Ashish Vaswani、Noam Shazeer、Niki Parmar、Jakob Uszkoreit、Llion Jones、Aidan N. Gomez、Lukasz Kaiser 和 Illia Polosukhin。在 Transformer 诞生之前，序列到序列（Seq2Seq）模型主要依赖于循环神经网络（RNN）和其变体（如 LSTM、GRU），这些模型在处理长序列时存在梯度消失、并行化困难等问题。

Transformer 模型主要解决了以下几个关键问题：
 并行化问题：传统 RNN 模型的顺序处理特性使其难以并行化，而 Transformer 通过自注意力机制实现了完全并行化；
 长程依赖问题：RNN 在处理长序列时存在梯度消失问题，Transformer 通过直接建模任意两个位置之间的关系来解决这一问题；
 计算效率问题：相比于 RNN 的顺序计算，Transformer 的矩阵运算可以充分利用现代 GPU 的并行计算能力；
 表征能力问题：自注意力机制使模型能够更好地捕捉序列中的全局依赖关系。

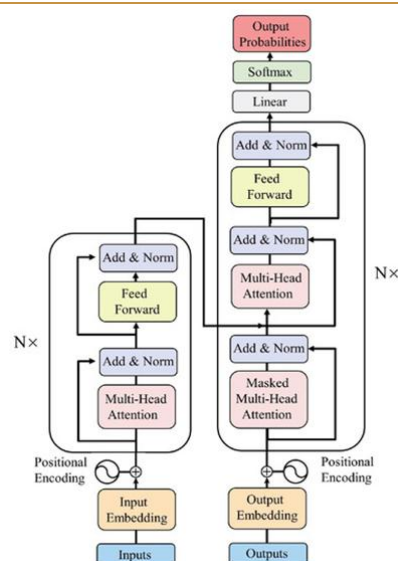
（1）模型整体结构

Transformer 采用编码器-解码器（Encoder-Decoder）架构，由功能互补的两部分组成：

编码器（Encoder）：由 N 个（原论文 $N=6$ ）相同层堆叠而成，负责提取输入序列的深层特征表示。每层包含两个核心子层：多头自注意力机制和前馈神经网络（FFN）。每个子层均采用残差连接（Residual Connection）和层归一化（Layer Normalization），确保训练稳定性。

解码器（Decoder）：同样由 N 个相同层构成，在生成输出序列时，通过掩蔽自注意力确保当前位置仅访问已生成标记。通过编码器-解码器注意力整合输入序列信息。解码器的独特设计使其适用于自回归生成任务，如文本生成。

图 1：Transformer 网络结构



数据来源：西南证券整理

(2) 注意力机制

Transformer 的核心是注意力机制 (Attention Mechanism)，特别是自注意力 (Self-Attention) 机制。注意力机制的本质是计算查询 (Query)、键 (Key) 和值 (Value) 之间的相关性：

$$Q = XW_Q, K = XW_K, V = XW_V$$

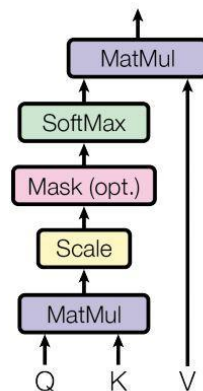
其中 W_Q, W_K, W_V 分别代表查询 (Query)、键 (Key)、值 (Value) 矩阵对应的线性变换矩阵。此外，通过点积 Q, K 计算注意力得分，进而得到自注意力值 Z ：

$$Z = \text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

其中 d_k 为键向量的维度，用于防止梯度消失， softmax 函数将一组数值转换为概率分布 (总和为 1)，放大高分值的相对权重，从而实现注意力得分的归一化。

图 2：注意力机制

Scaled Dot-Product Attention



数据来源：西南证券整理

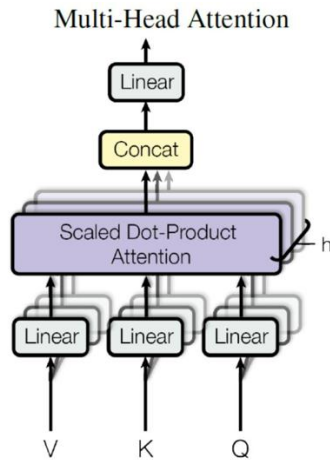
(3) 多头注意力机制

除了上述提及的简单的自注意力计算外，有时也会使用多头注意力机制，即在构建 Transformer 模型时也会将 Q, K, V 分割为多个子空间并行计算，从而得到数个注意力头 z_i 并拼接，最后通过线性变换矩阵，将拼接结果变换为一个融合了所有注意力头信息的矩阵 z ：

$$z = \text{Multihead}(Q, K, V) = \text{Concat}(z_1, \dots, z_n)W_O$$

其中，每个注意力头 z_i 独立学习不同子空间的特征表示， W_O 表示融合后的线性变换矩阵。

图 3：多头注意力示意



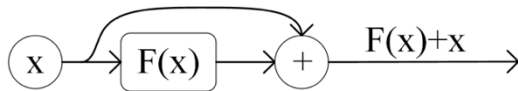
数据来源：西南证券整理

(4) 前馈网络与残差连接

自注意力输出经过前馈网络（含ReLU激活）和层归一化（LayerNorm），并通过残差连接缓解梯度消失问题。算式如下，其中， $Sublayer(x)$ 指当前子层（自注意力层或前馈层）输出（如前述提及的 z ）

$$Output = LayerNorm(x + Sublayer(x))$$

图 4：残差连接示意



数据来源：西南证券整理

前馈网络（Feed-Forward Network, FFN）层比较简单，是一个两层的全连接层，第一层的激活函数为 ReLU，第二层不使用激活函数，对应的公式如下。

$$FNN(X) = \max(0, WX + b_1) W_2 + b_2$$

其中 X 是输入，前馈网络最终得到的输出矩阵的维度与 X 一致。

上述提及的ReLU函数可以通过将负输入置零并保留正输入，提升网络的稀疏性与泛化能力，并促进计算高效性，其算式如下：

$$ReLU(x) = \max(0, x)$$

表 1：Transformer 模型各结构功能及创新

模块	功能	关键创新
编码器	提取输入序列语义特征	多头自注意力实现上下文编码
解码器	基于编码特征生成目标序列	掩码注意力防止信息泄漏
编码器-解码器注意力	对同源语言与目标语言关键词	实现跨序列软对齐

数据来源：西南证券整理

解码器通过掩码自注意力确保当前位置只能访问先前位置信息，维持自回归特性；同时通过编码器-解码器注意力建立输入与输出的动态关联。

1.2 Transformer 与传统序列模型对比

(1) 模型架构对比

传统序列模型特点：

RNN（循环神经网络）：基本的循环结构，存在严重的梯度消失问题。

LSTM（长短期记忆网络）：通过门控机制缓解梯度消失，但计算复杂度较高。

GRU（门控循环单元）：简化的 LSTM，参数较少但性能接近。

Transformer 特点：完全基于注意力机制，无循环结构；编码器-解码器架构；多头注意力并行处理。

(2) 计算复杂度对比

表 2：不同模型计算复杂度性能特点对比

模型类型	时间复杂度	空间复杂度	并行化能力	最小操作数
RNN	$O(nd)$	$O(d)$	不支持	$O(n)$
LSTM	$O(nd)$	$O(d)$	不支持	$O(n)$
GRU	$O(nd)$	$O(d)$	不支持	$O(n)$
Transformer	$O(n^2d)$	$O(n^2)$	完全支持	$O(1)$

数据来源：西南证券整理

其中： n 为序列长度； d 为隐藏层维度

时间复杂度：RNN 系列对序列长度呈线性关系，Transformer 对序列长度呈二次关系。

空间复杂度：RNN 系列空间需求固定，Transformer 需要存储注意力矩阵。

并行化：RNN 系列必须顺序计算，Transformer 可以完全并行。

(3) 梯度传播特性与长程依赖建模能力

表 3：不同模型梯度传播特性与长程依赖建模能力

模型	梯度路径长度	100 步后梯度保留率	1000 步后梯度保留率
RNN	$O(n)$	<0.1%	≈0%
LSTM	$O(\log n)$	12.7%	0.8%
Transformer	$O(1)$	>99.9%	>99.5%

数据来源：西南证券整理

RNN 系列：信息需要逐步传递，容易丢失；梯度在反向传播中指数衰减；实际有效长度受限（通常<100）。

LSTM/GRU：通过门控机制改善，但仍有限制；能处理稍长序列（数百个时间步）；计算复杂度增加。

Transformer: 任意两个位置直接建模；理论上无长程依赖限制；受限于计算资源和序列长度限制。Transformer 通过残差连接和层归一化实现近乎无损的梯度传播，基本解决 RNN 家族的梯度消失问题。

1.3 Transformer 主要优势

(1) 并行化能力

Transformer 的自注意力机制通过矩阵乘法实现全局关联计算，该计算可分解为独立子矩阵运算，天然适配 GPU/TPU 的大规模并行架构。

(2) 长程依赖建模

通过直接计算任意两个位置之间的注意力权重，Transformer 能够有效捕捉长程依赖关系。传统 RNN/LSTM 的路径长度为 $O(n)$ ，Transformer 降低至 $O(1)$ ：任意两位置直接关联，彻底解决梯度消失/爆炸问题。

(3) 可解释性

注意力权重提供了模型决策过程的可视化，增强了模型的可解释性。

(4) 迁移学习能力

在计算机视觉领域，Vision Transformer (ViT) 将图像分割为 16×16 的 Patch 序列，通过位置编码保留空间信息，利用 Transformer 编码器提取全局特征。在 ImageNet-21k (含 1,400 万图像) 预训练预训练后，模型学习到通用视觉表征。在 CIFAR-100 数据集上微调后，ViT-L/16 模型达到 94.3% 准确率，显著超越相同条件下 ResNet152 的 93.3%。ViT 证明了纯 Transformer 架构在 CV 任务中的通用性。(Dosovitskiy et al. (2020). An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale. ICLR 2021.)

在生物医学这一高度专业化的领域，Transformer 模型同样展现出了强大的适应性。BioBERT 作为专门针对生物医学文本优化的 BERT 变体，通过在 PubMed 大规模生物医学语料库上进行预训练，成功地学习到了生物医学领域的专业知识表示。当将这种预训练模型应用于生物医学命名实体识别任务时，相比于通用语言模型，BioBERT 的 F1 值显著提升了 8.2%。这一显著的性能提升表明，领域特定的预训练能够让 Transformer 模型更好地理解 and 处理专业领域的复杂语义关系，从而在实际应用中发挥更大的价值。(Lee, J., Yoon, W., Kim, S., et al. (2020). BioBERT: a pre-trained biomedical language representation model for biomedical text mining. Bioinformatics, 36(4), 1234-1240.)

在代码生成和程序合成领域，Codex，作为 GPT-3 的专门变体，专注于理解和生成程序代码。其最令人印象深刻的特性在于极强的跨编程语言迁移能力——当需要适应全新的编程语言时，Codex 仅需要少量的示例样本（通常只需要数百个样本）进行微调，就能够掌握新语言的语法规则和编程范式。这种 few-shot 学习能力展现了 Transformer 模型对代码结构和逻辑的深层理解。(Chen, M., Tworek, J., Jun, H., et al. (2021). Evaluating large language models trained on code. arXiv:2107.03374.)

1.4 传统 Transformer 模型局限性及可优化点

(1) 计算复杂度问题

自注意力的二次复杂度：时间复杂度： $O(n^2d)$ ，其中 n 是序列长度， d 是模型维度；空间复杂度： $O(n^2)$ ，需要存储注意力矩阵。对于长序列，这种二次复杂度会导致计算和内存消耗急剧增加。

(2) 数据和计算资源需求

高资源需求：需要大量的训练数据才能达到最佳性能。训练需要大量的计算资源和时间并且模型参数量巨大，部署成本高。

(3) 任务特定缺陷：在不同应用场景中暴露特有局限

时间序列预测：通道不独立：多变量强行拼接破坏物理意义（如温度与股价尺度差异）；点式注意力不匹配：时间局部连续性未被有效利用，PatchTST 证明分块（Patch）更合理。

数学推理：符号运算和逻辑推导能力弱，需注入外部模块；思维链（Chain of Thought）依赖注意力精细权重分配，部分高效模型在此表现不佳。

本报告致力于深入探究并显著提升 Transformer 深度学习模型在 A 股市场量化选股策略中的应用效能。尽管 Transformer 凭借其强大的序列建模能力在自然语言处理等领域取得革命性成功，但其在金融时序数据（特别是股票收益预测）上的直接迁移应用仍面临独特挑战，尽管已有研究证明 Transformer 模型在 A 股市场具备较好的选股效果，但其存在进一步优化的空间。

研究伊始，我们将严谨构建并测试一个标准 Transformer 模型作为性能基准。通过评估未经优化的 Transformer 模型在选股任务中的原始表现，揭示其在捕捉金融市场复杂动态（如非平稳性、长短期依赖、噪声干扰）方面存在的固有局限，为后续优化提供明确的方向与对比基线。基于西南金工 2025 年 3 月发布的报告《PINN 信息约束与时序截面双流网络模型》，本文训练模型采取的时序特征如下：

表 4：Transformer 模型量价时序特征列表

序号	特征中文名称	特征助记符	特征计算方法
1	量价背离度	pvo	T 收盘价减去收盘价的 5 日指数移动平均值后，除以成交量的 20 日移动平均值
2	压力支撑效率	sse	过去 20 日最高收盘价减去过去 20 日最低收盘价后，除以成交量的 20 日移动平均值
3	流动性冲击系数	liq	T 成交量与昨日成交量差值的绝对值，除以过去 20 日内成交量序列的标准差
4	波动率偏度	skew	T 日收益率过去 20 日内的三阶标准化矩
5	隔夜跳空强度	gap	T 开盘价减去昨日收盘价，再除以 T-1 收盘价
6	波动率期限结构	vts	过去 10 日内日收益率序列的标准差除以过去 30 日内日收益率序列的标准差
7	量能聚集度	vcil	T 成交量除以过去 20 日内成交量的移动平均值
8	收益率波动比	rvr	T 收益率除以过去 20 日序列日收益率标准差
9	筹码松动度	csl	T 收盘价减去过去 60 日收盘价的移动平均值后，除过去 60 日收盘价的移动平均值
10	反转效应强度	rev	过去 20 日序列中，当日收益率与昨日日收益率的相关系数
11	波动率聚集度	vclb	过去 5 日内日收益率序列的标准差除以过去 20 日内日收益率序列的标准差
12	收益分布峰度	kurtosis	日收益率过去 20 日内的四阶标准化矩

序号	特征中文名称	特征助记符	特征计算方法
13	流动性冲击持续性	lip	进三日流动性冲击系数加权求和
14	市场宽度指标	bread	当日上涨股票数减去当日下跌股票数，除以总股票数

数据来源：西南证券整理

模型训练方式等其他模型设定如下：

- 1) **训练数据**：所有个股过去 10 年内的时序特征，做 5 天采样，预测标签为未来一个月 (20 个交易日) 的累计收益；
- 2) **训练集验证集比例**：80%：20%；
- 3) **时序尺度选择**：40 天；
- 4) **数据处理**：时序标准化（与最后一个时间步比值），截面标准化（均值标准差）；
- 5) **模型训练及调仓频次**：一年更新，月频调仓；
- 6) **模型超参数**：batch=截面个股数，lr=1e-4，损失函数 1-IC。

自 2019 年 1 月至 2025 年 7 月，基础 Transformer 因子 IC(以 RankIC 表征)为 9.42%，多头组合年化收益率为 28.70%，月均单边换手率 0.88X，整体来看，仅靠量价时序特征训练出的基础 Transformer 模型具备一定的选股效果。(IC 测试中“***”、“**”、“*”分别代表参数在 1%、5%、10% 的显著性水平下显著，下同，不再特殊强调。)

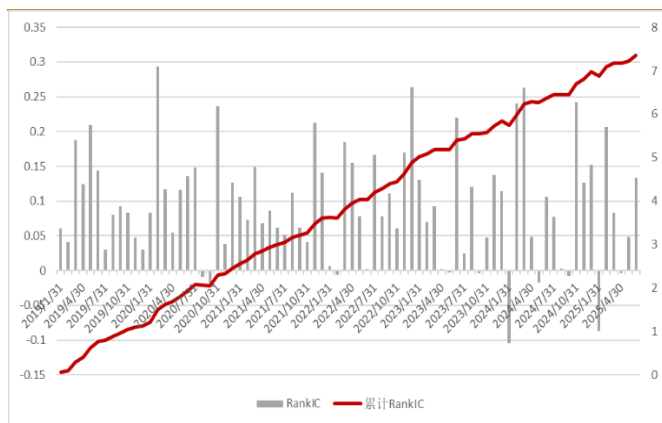
表 5：基础 Transformer 回测结果

	IC	ICIR	Top10%组合年化收益率	年化波动	信息比率	最大回撤率	单边月均换手率
基础 Transformer	0.0942***	4.03	28.70%	24.81%	1.16	39.43%	0.87X

数据来源：wind、恒生聚源、ifind、西南证券整理，截至 2025 年 7 月 31 日

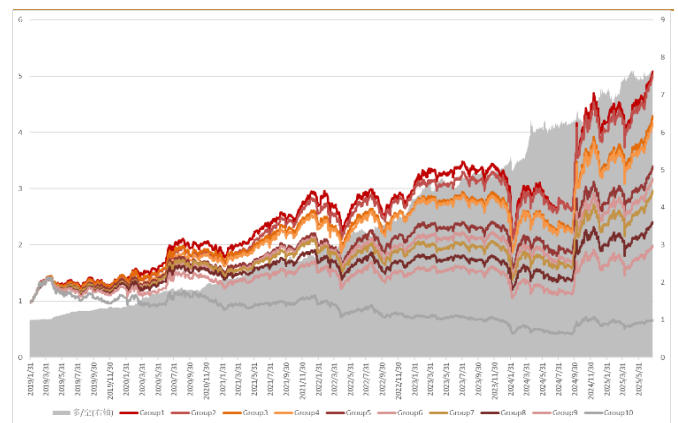
从分组角度来看，多头组合（组 1）年化收益率 28.70%，最大回撤 39.42%；空头组合（组 10）年化收益率-6.17%，最大回撤 70.95%；多空组合（组 1/组 10）年化收益率 37.16%，最大回撤 9.30%，具备一定单调性。但整体来看，因子多头组合表现并不显著，且除空头组合外，组一-组二组合分化程度不高，基础 Transformer 模型及量价时序信息在因子挖掘效果层面更多聚焦于空头，具备进一步提升空间。

图 5：基础 Transformer 因子 IC



数据来源：wind、ifind、西南证券整理，截至 2025 年 7 月 31 日

图 6：基础 Transformer 因子分组-绝对净值

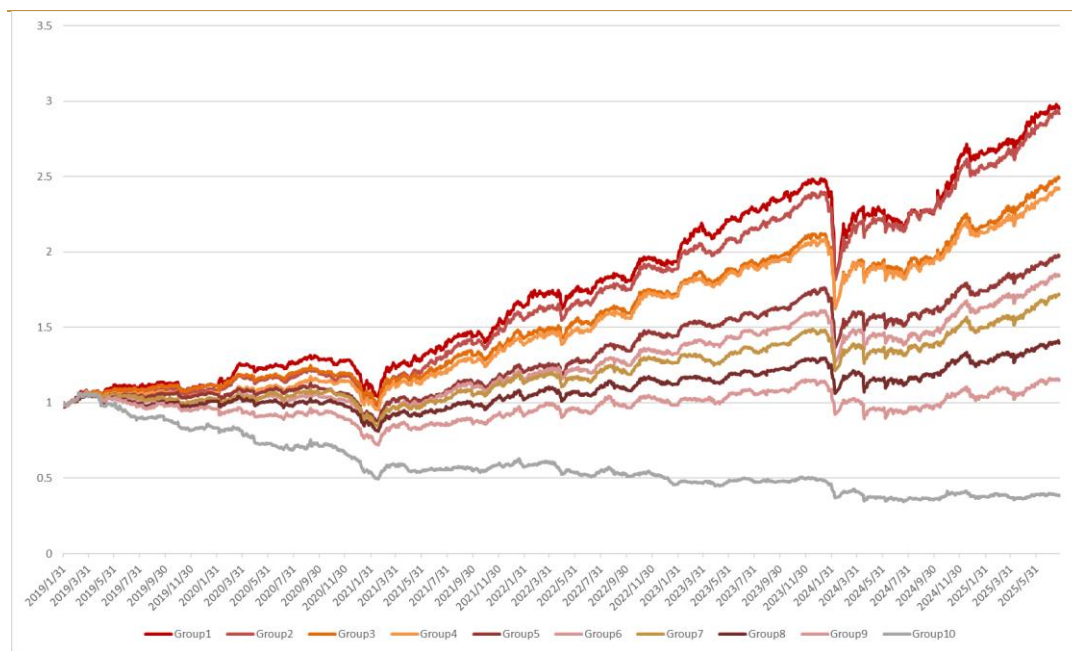


数据来源：wind、ifind、西南证券整理，截至 2025 年 7 月 31 日

表 6：基础 Transformer 因子 10 分组表现

	年化收益率	年化波动	信息比率	最大回撤率
Group1	28.70%	24.81%	1.16	39.43%
Group2	28.47%	24.22%	1.18	36.93%
Group3	25.37%	24.27%	1.05	36.77%
Group4	24.77%	24.33%	1.02	35.77%
Group5	20.92%	24.56%	0.85	34.96%
Group6	19.64%	24.62%	0.80	35.20%
Group7	18.33%	24.83%	0.74	34.81%
Group8	14.62%	25.19%	0.58	36.99%
Group9	11.20%	25.76%	0.43	38.63%
Group10	-6.17%	27.48%	-0.22	70.95%
多/空	37.16%	11.82%	3.15	9.30%

数据来源：wind、恒生聚源、ifind、西南证券整理，截至 2025 年 7 月 31 日

图 7：基础 TransFormer 因子分组测试-相对净值


数据来源：wind、ifind、西南证券整理，截至 2025 年 7 月 31 日

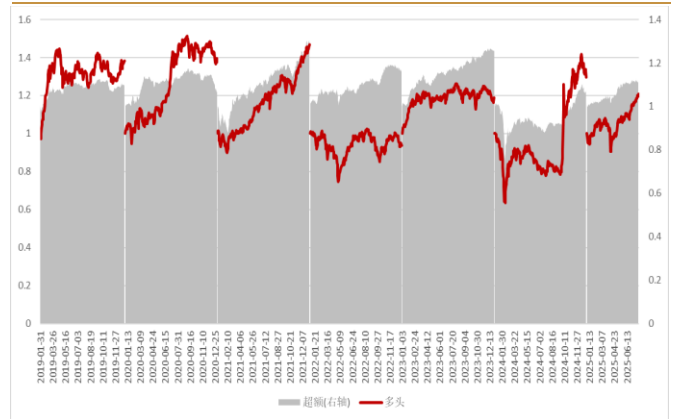
从因子分组的相对净值表现来看，相对万得全 A，Transformer 因子在 2024 年 1 月底出现了较大的超额回撤，但在 2024 年 9 月底的上涨行情（0924 行情）内并未出现较为明显的超额回撤。

分年度来看，多头组合自 2019 年至 2025 年 7 月，7 年均跑赢万得全 A 指数，其中 2021 年相对万得全 A 年化超额 30.55%，超额最大回撤 12.65%；2023 全年相对万得全 A 年化超额收益率 25.44%，超额最大回撤仅 4.87%。此外，因子今年以来超额表现较好，2025 年初至 7 月 31 日，累计超额收益率 11.03%，超额最大回撤 3.17%。受限于 2024 年两端极端行情，全年多头组合相对万得全 A 超额进 6.37%，超额回撤 26.26%。

表 7：基础 Transformer 因子多头分年度超额表现

年份	年化收益率	年化波动率	信息比率	最大回撤率
2019	9.34%	8.30%	1.13	5.13%
2020	6.41%	9.04%	0.71	9.37%
2021	30.55%	14.37%	2.13	12.65%
2022	16.20%	9.94%	1.63	7.13%
2023	25.44%	7.38%	3.45	4.87%
2024	6.37%	18.73%	0.34	26.26%
202507	11.03%	7.67%	1.44	3.17%

数据来源：wind、ifind、西南证券整理，截至 2025 年 7 月 31 日

图 8：基础 Transformer 因子多头分年度净值


数据来源：wind、ifind、西南证券整理，截至 2025 年 7 月 31 日

传统 Transformer 模型具备一定的选股效果，但模型整体选股效果和模型稳定性具备较大的优化空间，且在关键市场阶段（2024 年 1 月量化危机、2024 年 0924 行情）仍会出现较为明显的超额回撤。

本报告的提出了一类基于 Transformer 模型自适应的解决方案（Dynamic Adaptive Fusion Attention Transformer, DAFAT），主要工作点在于系统性地提出并实现针对 Transformer 架构三大关键组件的深度优化：

位置编码自适应：针对金融时间序列的强时序依赖性与周期性（如季节性效应、事件驱动），我们针对标准正弦/余弦位置编码，引入动态时变位置编码机制，旨在精确地捕捉因子在时间轴上的相对重要性演变与市场状态影响的持续效应，提升模型对时序结构的感知能力。

注意力机制自适应：传统全局注意力在高噪声、高维金融数据上易引入冗余信息并导致过拟合。我们设计稀疏化与局部敏感的注意力机制，引导模型聚焦于当前预测最具信息量的少数关键因子及历史片段，旨在提升计算效率，抑制噪声干扰。

信息融合自适应：为充分利用不同来源（量价、基本面等）和不同频率（日、周、月、季）因子，本文设计多尺度信息融合框架，其多尺度信息融合框架包含微观尺度、中观尺度以及宏观尺度，并引入跨尺度注意力门控融合机制以及原始量价特征残差融合，旨在超越简单的拼接或加权，实现跨因子、跨周期的深度非线性交互与自适应信息整合，增强模型对复杂市场微观结构的刻画能力。

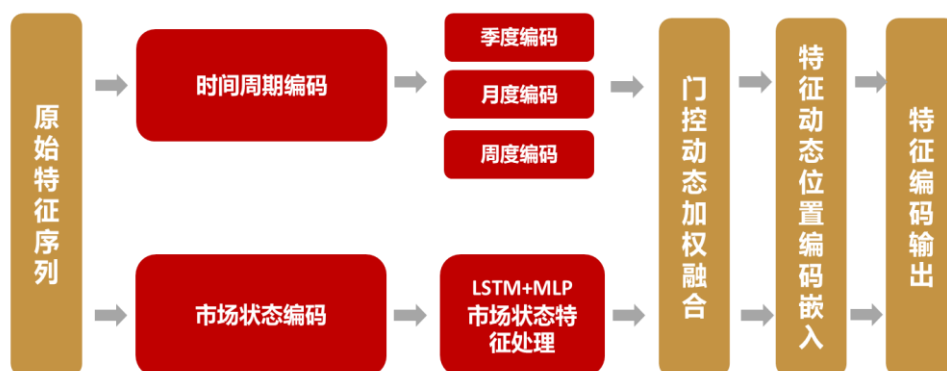
2 位置编码自适应：动态位置编码解决方案

在传统 Transformer 选股模型中，位置编码机制存在三重核心缺陷，可能会严重制约模型在金融时序数据上的表现。

(1) 周期性失配问题：固定正弦/余弦编码无法适配金融市场的独特周期（例如特定的财报季、年末效应、政策窗口期）。传统编码与金融季度周期的相关性较低，有可能导致模型无法准确捕捉各类事件驱动的行市；

(2) 状态盲区问题：静态编码忽视市场状态转换对因子有效性的根本影响。如图 1 所示，动量因子在牛熊市中的有效性差异达 80%，但传统模型对此完全失效，造成熊市超额回撤。

图 9：动态位置编码自适应：时间周期编码+市场状态编码



数据来源：西南证券整理

基于上述问题，本文提出了双通道门控动态位置编码（Double-Gate DPE）以实现 Transformer 模型中的位置编码自适应。其中双通道门控动态位置编码分别包括时间周期编码以及市场状态编码，两类编码信息分别由不同的深度学习神经网络模型处理后，经有门控融合网络对两类动态位置编码进行动态加权求和，最终输出新的动态位置编码方式。

动态位置编码之时间周期编码：为了更精确的刻画时间日历效应对个股收益预测的影响，改变以往固定正弦/余弦编码无法适配金融市场的独特周期的问题，本文引入周度、月度以及季度时间周期信息作为时间位置编码信息。其中时间周期位置编码信息处理步骤如下：

- 1) 季度编码：0-3，表示四个季度；月度编码：0-11 表示 12 个月；周度编码：0-4 表示一周五天。以上编码均为整数编码；
- 2) 整数编码周期化调整：引入正余弦函数对上述编码进行处理，假设编码周期为 n ，则周期化调整公式为：

$$\text{sincomponent}(x) = \sin(2\pi x/n)$$

$$\text{coscomponent}(x) = \cos(2\pi x/n)$$

- 3) 时间编码嵌入：合并季度、月度、周度周期化调整后的信息并引入全连接层嵌入。

动态位置编码之市场状态编码：引入市场波动率（市场情绪）、行业轮动速度（热点切换）以及市场流动性（资金活跃度）三维状态指标来表征市场状态，并利用时序模型 LSTM+MLP 的深度学习网络结构进行市场状态信息的处理，并最终输出多维市场状态信息。其中市场波动率、行业轮动速度以及市场流动性的特征表示如下：

- 1) **市场波动率：**全市场个股 20 日收益率标准差中位数；
- 2) **行业轮动速度：**行业 20 日累计收益标准差（申万一级行业分类）；
- 3) **市场流动性：**全市场换手率。

在构建时间周期位置编码和市场状态位置编码后，引入门控融合机制，即引入可学习线性层，动态学习两类动态位置编码的合成权重，并合成两类动态位置编码。动态位置编码作为 Transformer 模型自适应改造的第一部分（特征处理部分），主要采取时间周期编码自适应和市场状态自适应，并最后并引入可学习参数的门控融合层对两个不同时间尺度的编码信息进行加权求和，将最后加权求和后的位置编码信息作为最终的 Transformer 的位置编码信息。

2.1 动态位置编码自适应 Transformer 模型回测结果

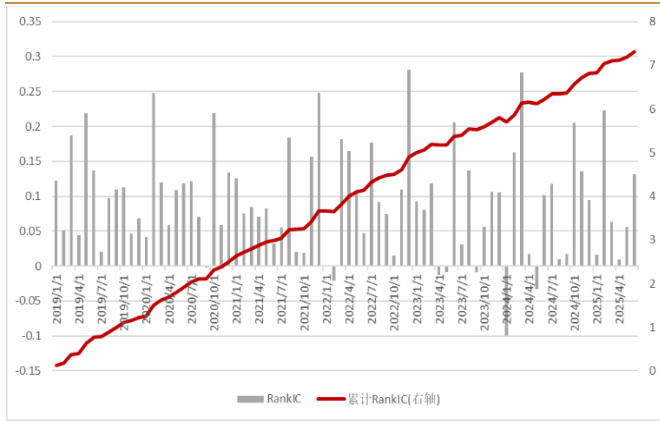
本节中仅测试加入动态位置编码自适应的 Transformer 模型的选股因子效果（DPE_Transformer, Dynamic Positional Encoding Transformer）。自 2019 年 1 月至 2025 年 7 月，DPE_Transformer 因子 IC 均值为 9.54%，多头组合年化收益率 30.60%，月均单边换手率 0.86X。整体 IC 及多头组合表现优于基础 Transformer 模型，换手率小幅低于基础 Transformer 模型。

表 8：DPE_Transformer 回测结果

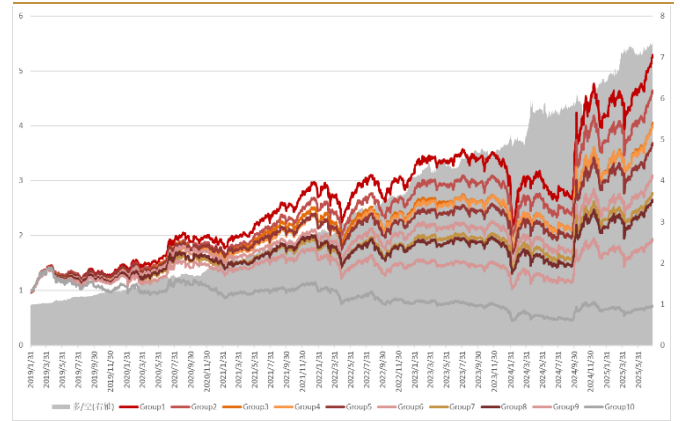
	IC	ICIR	Top10%组合年化收益率	年化波动率	信息比率	最大回撤率	单边月均换手率
DPE_Transformer	0.0954***	4.44	30.60%	24.97%	1.23	39.41%	0.86X

数据来源：wind、恒生聚源、ifind、西南证券整理，截至 2025 年 7 月 31 日

从分组角度来看，多头组合（组 1）年化收益率 30.60%，最大回撤 39.41%；空头组合（组 10）年化收益率-5.02%，最大回撤 68.41%；多空组合（组 1/组 10）年化收益率 37.23%，最大回撤 8.69%，整体具备一定单调性。且相对于传统 Transformer 模型分组测试中组一与组二分化不明显的现象，DPE_Transformer 模型因子单调性具备一定程度的提升。

图 10: DPE_Transformer 因子 IC


数据来源: wind、ifind、西南证券整理, 截至 2025 年 7 月 31 日

图 11: DPE_Transformer 因子分组-绝对净值


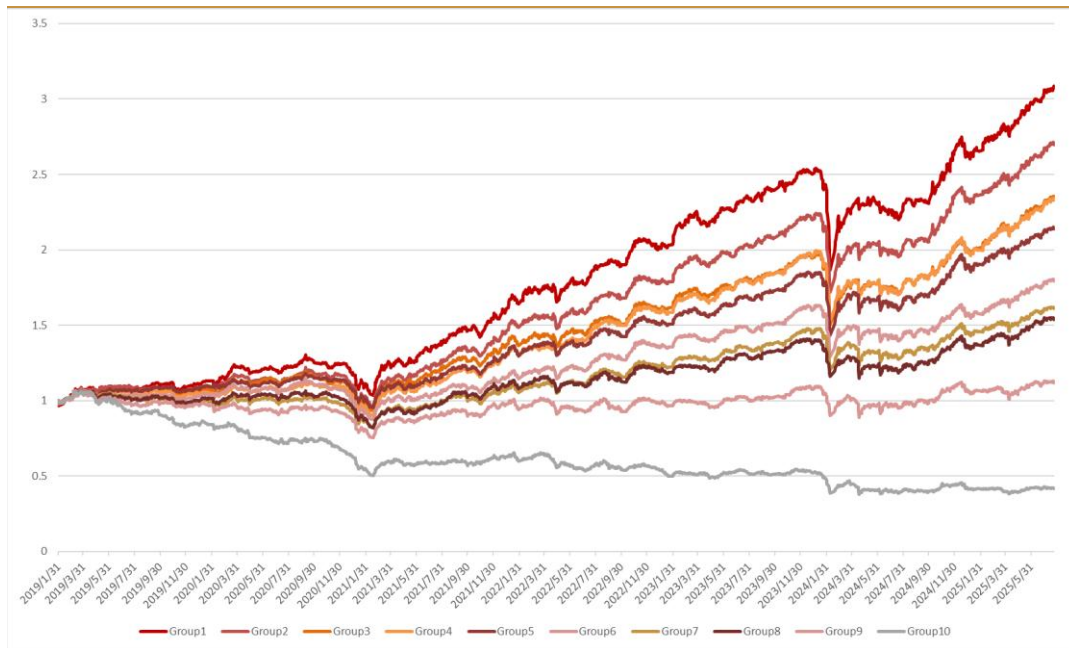
数据来源: wind、ifind、西南证券整理, 截至 2025 年 7 月 31 日

表 9: DPE_Transformer 因子 10 分组表现

	年化收益率	年化波动	信息比率	最大回撤率
Group1	30.60%	24.97%	1.23	39.41%
Group2	26.89%	24.28%	1.11	36.55%
Group3	24.24%	24.41%	0.99	37.11%
Group4	24.07%	24.36%	0.99	35.71%
Group5	22.43%	24.56%	0.91	35.73%
Group6	19.18%	24.61%	0.78	34.83%
Group7	17.18%	24.80%	0.69	33.48%
Group8	16.35%	25.10%	0.65	35.06%
Group9	10.76%	25.69%	0.42	42.41%
Group10	-5.02%	27.25%	-0.18	68.41%
多空	37.23%	11.52%	3.14	8.70%

数据来源: wind、恒生聚源、ifind、西南证券整理, 截至 2025 年 7 月 31 日

从因子分组的相对净值表现来看, 相对万得全 A, 多尺度 Transformer 因子在 2024 年 1 月底出现了较大的超额回撤, 但在 2024 年 9 月底的上涨行情 (0924 行情) 内并未出现较为明显的超额回撤。

图 12: DPE_Transformer 因子分组测试-相对净值


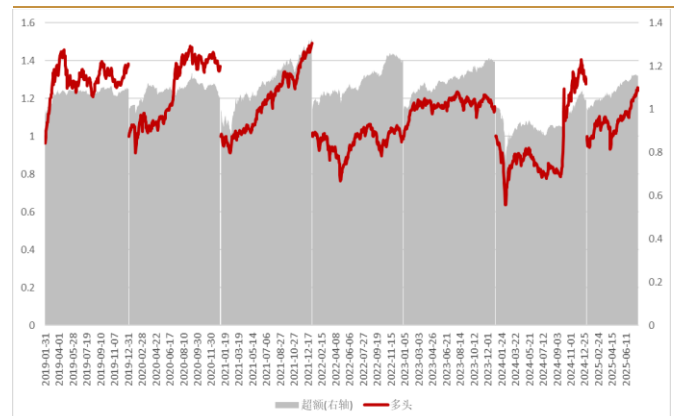
数据来源: wind、ifind、西南证券整理, 截至2025年7月31日

分年度来看, 多头组合自 2019 年至 2025 年 7 月, 7 年均跑赢万得全 A 指数, 其中 2021 年相对万得全 A 年化超额 32.54%, 超额最大回撤 11.51%; 2022、2023 全年相对万得全 A 年化超额收益率均超过 20%, 超额最大回撤分别仅 6.53%、4.59%。此外, 因子今年以来超额表现较好, 2025 年初至 7 月 31 日, 累计超额收益率 16.02%, 超额最大回撤 2.81%。

表 10: DPE_Transformer 因子多头分年度超额表现

年份	年化收益率	年化波动率	信息比率	最大回撤率
2019	9.36%	8.17%	1.15	4.33%
2020	4.90%	8.68%	0.56	10.23%
2021	32.54%	13.92%	2.34	11.51%
2022	22.24%	10.38%	2.14	6.53%
2023	22.12%	7.22%	3.06	4.59%
2024	4.48%	18.00%	0.25	25.88%
202507	16.02%	7.79%	1.93	2.81%

数据来源: wind、ifind、西南证券整理, 截至2025年7月31日

图 13: DPE_Transformer 因子多头分年度净值


数据来源: wind、ifind、西南证券整理, 截至2025年7月31日

2024 年 DPE_Transformer 多头组合相对于万得全 A 超额仅 4.48%, 在加入时间周期编码与市场状态编码组成的动态位置编码信息后, DPE_Transformer 模型在 2024 年 1 月末及 2024 年“0924”行情下相对万得全 A 仍然会产生较大幅度的回撤。

3 注意力自适应：稀疏注意力解决方案

在金融时序预测领域，传统 Transformer 模型的全连接注意力机制存在一定的局限性，亟需针对性优化。本文深入研究发现，标准注意力机制在应用于股票数据时可能会面临如下三重挑战：

首先，全连接注意力计算复杂度较高。其 $O(n^2)$ 计算复杂度难以支撑全市场高达 5000 只股票的高效计算，尤其针对多个交易日的长序列场景下，内存消耗较高，计算效率较低，将会制约模型的可扩展性；**其次，全连接结构对噪声高度敏感。**当市场状态快速切换的时候，全连接注意力机制极有可能产生较多的无效注意力权重，这将会干扰模型判断。

针对这些痛点，本文提出三重稀疏化注意力自适应机制，该机制可以在提升计算效率的同时，加大注意力机制的局部针对性聚焦模式，减少对噪声时段的关注。波动率门控作为第一道过滤器，通过可学习的波动率预测网络，自动识别并屏蔽波动率低于一定程度的低信息时段，同时确保每行至少保留一定的关键连接。局部注意力窗口则聚焦小窗口的短期模式，这与金融数据的动量周期完美契合，通过限制注意力范围强化对突破形态、反转信号等关键模式的捕捉。最后，Top-k 稀疏选择动态保留每行前 30% 的强相关连接，减少大部分计算量，使模型能实时处理全市场数据。

(1) 波动率门控：首先计算各时间点的过去 N 日的收益率波动率 σ_t ($N=20$)，然后根据波动率生成二值掩码（当波动率小于一定阈值的情况下掩码设为 0），并在每行保留波动率前 30% 的连接以防止过度稀疏。其中二值掩码计算公式如下

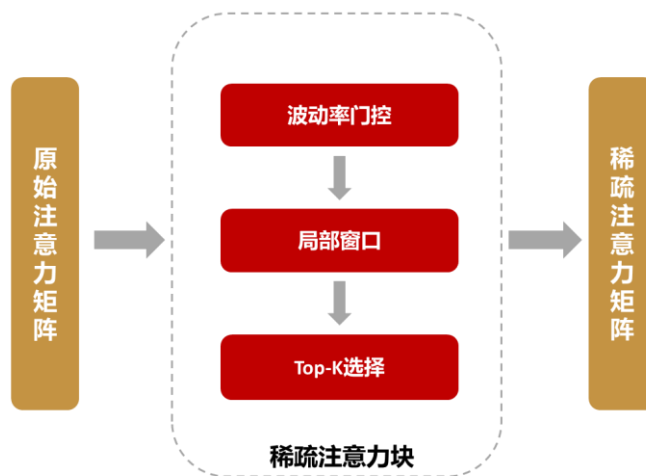
$$\text{Mask}_t = \begin{cases} 0, & \sigma_t < 0.1 \\ 1, & \text{otherwise} \end{cases}$$

波动率门控主要目的为根据波动率来过滤注意力连接。具体来说，它识别并屏蔽低波动率时段的注意力连接（通常低波动时段信息量较少），尽可能保留高波动率时段的连接（高波动时段信息丰富）。操作上主要通过对每个时间步计算波动率，如果波动率低于设定的阈值，则将该时间步对应的注意力分数设置为一个极小的值（如负无穷），这样在后续的 softmax 中这些位置的概率接近 0。

(2) 局部注意力窗口：该步骤的主要目的为将依赖范围约束为相对短期的模式，即限制每个查询（query）只能关注一个局部窗口内的键（key）。这个窗口通常以查询位置为中心，覆盖前后一定的时间步。该步骤可减少计算量，并同时强化模型对局部模式的捕捉能力。模型仍会接受完整的量价时序信息（seqlen=60），但每个时间点只关注其前后一定长度的局部窗口（注意力机制被限制在局部）。

具体操作上，创建一个带状掩码（band mask），只允许在窗口内的位置计算注意力分数，窗口外的位置被屏蔽。

(3) Top-k 稀疏选择：在波动率门控和局部窗口约束后，进一步对每行注意力权重实施 Top K 选择，即对每个查询的注意力分数向量，选择分数最高的 K 个位置。该操作可以实现双重目标：从信息论角度，聚焦信息量最大的关键连接；从计算角度，可进一步降低计算复杂度。

图 14：注意力编码自适应：波动率门控+局部注意力窗口+Top-k 稀疏选择


数据来源：西南证券整理

在模型架构层面，我们将根据稀疏注意力模块构建稀疏注意力 Transformer 层，并代替原有的 Transformer 层。具体而言，在特征嵌入层之后，我们将其作为 Transformer 编码层的自注意力替代模块：在 QKV 投影计算初始注意力分数后，依次应用波动率门控、局部窗口和 Top-k 选择三重过滤，生成稀疏化注意力权重，再与价值向量相乘生成上下文表征。这种设计既保留了 Transformer 强大的序列建模能力，又通过先验信息引导注意力聚焦高价值信息，致力于使模型在噪声环境中保持稳健。

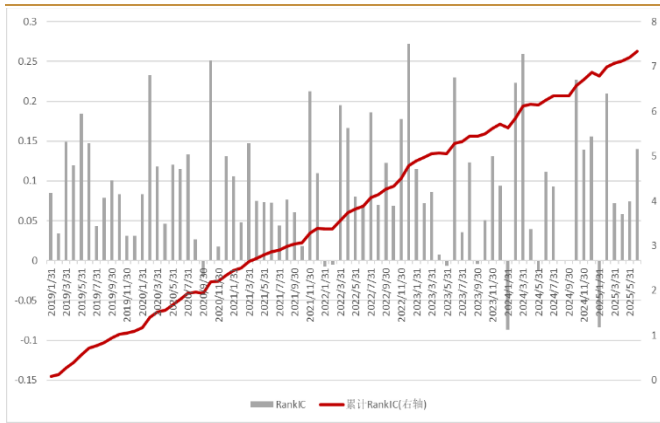
3.1 稀疏注意力自适应 Transformer 模型回测结果

本节中仅测试加入稀疏注意力自适应的 Transformer 模型的选股因子效果 (SA_Transformer, Sparse Attention Transformer)。自 2019 年 1 月至 2025 年 7 月，全 A 范围内 SA 模型因子月均 IC 为 9.98%，多头组合年化收益率为 29.64%，月均单边换手率 0.87X。整体来看，引入稀疏注意力机制的 Transformer 模型因子 IC 及多头组合表现略优于基础 Transformer 模型，因子换手相当。

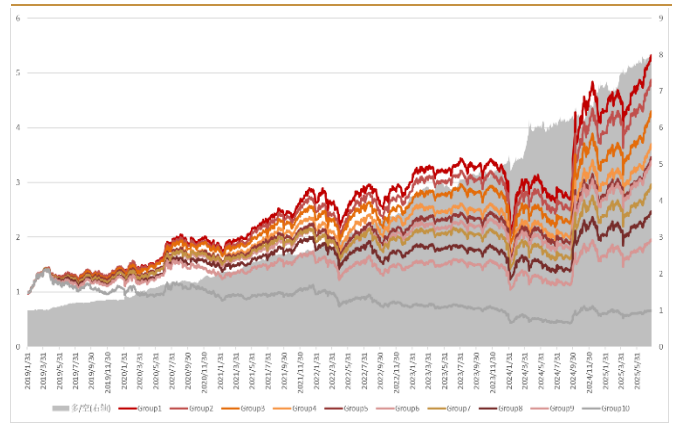
表 11：SA_Transformer 因子回测结果

	IC	ICIR	Top10%组合年化收益率	年化波动	信息比率	最大回撤率	单边月均换手率
SA_Transformer	0.0998***	4.24	29.64%	24.74%	1.20	38.62%	0.87X

数据来源：wind、恒生聚源、ifind、西南证券整理，截至 2025 年 7 月 31 日

图 15: SA_Transformer 因子 IC


数据来源: wind、i 西南证券整理, 截至 2025 年 7 月 31 日

图 16: SA_Transformer 因子分组-绝对净值


数据来源: wind、西南证券整理, 截至 2025 年 7 月 31 日

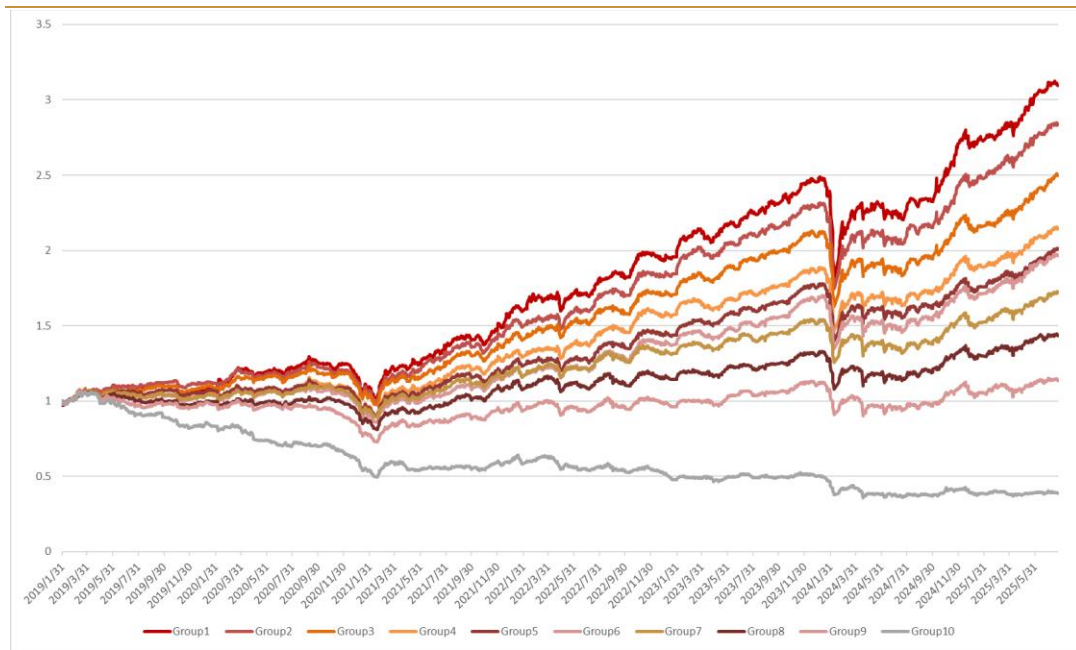
从分组角度来看, 多头组合 (组 1) 年化收益率 29.64%, 最大回撤 38.62%; 空头组合 (组 10) 年化收益率 -6.09%, 最大回撤 69.76%; 多空组合 (组 1 / 组 10) 年化收益率 38.03%, 最大回撤 8.61%, SA_Transformer 模型具备一定单调性。且相对于基础 Transformer 模型, SA_Transformer 在空头与多头均有提升。

表 12: SA_Transformer 因子分组表现

	年化收益率	年化波动	信息比率	最大回撤率
Group1	29.64%	24.74%	1.20	38.62%
Group2	27.87%	24.35%	1.14	37.57%
Group3	25.40%	24.33%	1.04	37.16%
Group4	22.50%	24.45%	0.92	36.25%
Group5	21.25%	24.53%	0.87	35.03%
Group6	20.84%	24.78%	0.84	34.71%
Group7	18.37%	24.99%	0.74	34.76%
Group8	15.06%	25.18%	0.60	38.75%
Group9	11.01%	25.61%	0.43	40.30%
Group10	-6.09%	27.00%	-0.23	69.76%
多/空	38.03%	11.64%	3.27	8.61%

数据来源: wind、恒生聚源、西南证券整理, 截至 2025 年 7 月 31 日

从因子分组的相对净值表现来看, 相对万得全 A, SA_Transformer 模型因子在 2024 年 1 月底同样出现了较大的超额回撤, 但随后回撤快速修复, 且在 2024 年 9 月底的上涨行情 (0924 行情) 内并未出现较为明显的超额回撤。

图 17: SA_Transformer 因子分组测试-相对净值


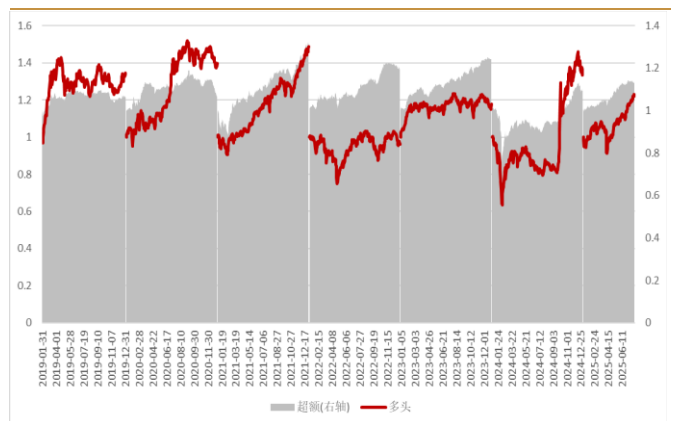
数据来源: wind、西南证券整理, 截至 2025 年 7 月 31 日

分年度来看, 多头组合自 2019 年至 2025 年 7 月, 7 年均跑赢万得全 A 指数, 其中 2021 年相对万得全 A 年化超额 32.24%, 超额最大回撤 11.98%; 2023 全年相对万得全 A 年化超额收益率 24.15%, 超额最大回撤仅 4.35%。此外, 因子今年以来超额表现较好, 2025 年初至 7 月 31 日, 累计超额收益率 12.93%, 超额最大回撤 2.99%。相较于基础 Transformer, SA_Transformer 模型在 2024 年表现有所提升, 其多头年化超额收益率为 9.46%, 超额回撤 26.18%, 优于基础 Transformer 模型多头超额收益率 6.37%, 超额回撤 26.26%。除 2024 年以外, SA_Transformer 于 2020、2021、2022、2024 及 2025 表现优于基础 Transformer。

表 13: SA_Transformer 因子多头分年度超额表现

年份	年化收益率	年化波动率	信息比率	最大回撤率
2019	6.14%	8.37%	0.73	5.66%
2020	6.52%	9.60%	0.68	10.57%
2021	32.24%	14.00%	2.30	11.98%
2022	19.48%	9.89%	1.97	7.01%
2023	24.15%	6.92%	3.49	4.35%
2024	9.46%	18.86%	0.50	26.18%
202507	12.93%	7.34%	1.76	2.99%

数据来源: wind、西南证券整理, 截至 2025 年 7 月 31 日

图 18: SA_Transformer 因子多头分年度净值


数据来源: wind、西南证券整理, 截至 2025 年 7 月 31 日

4 信息融合自适应：多尺度信息融合解决方案

在前文基础模型构建的特征选择过程中，本文除量价时序特征以外，选取了部分基本面特征并做日频化填充处理后，与量价时序特征合并，并放入 Transformer 模型进行训练。但是在金融时序建模中，我们面临一个根本性挑战：市场信息天然存在于不同时间尺度。量价数据以日频甚至更高频率更新，反映市场的瞬时波动和短期情绪；而基本面数据以季度频率发布，揭示企业的长期价值本质。传统方法简单地将基本面数据日频化处理后与量价数据拼接输入模型，导致三个核心问题：

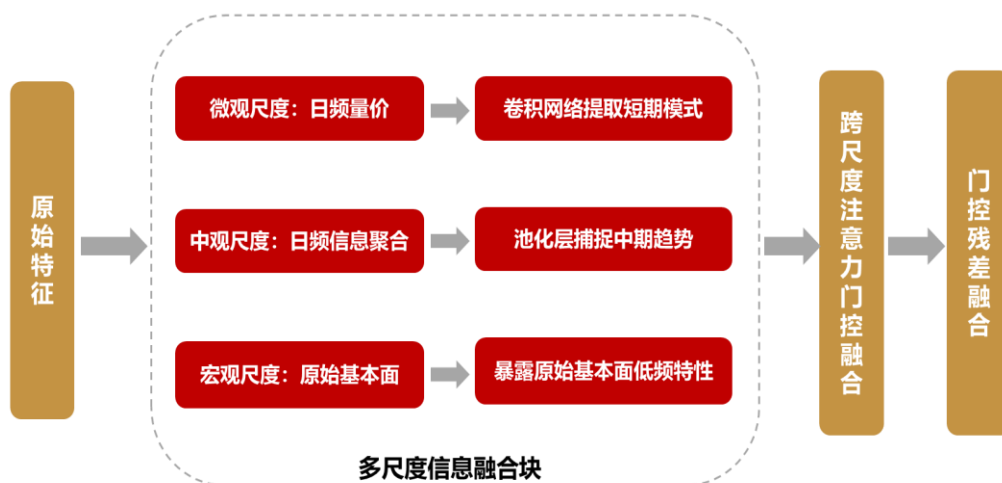
(1) 信息失真问题：基本面数据被强制填充为日频序列，极大部分数据点是通过插值生成的伪信号，不仅稀释了真实财报信息，还引入了大量噪声。

(2) 频率失配问题：量价特征的自相关滞后周期（约 1-5 天）与基本面特征（约 30-90 天）存在数量级差异。直接拼接导致模型混淆短期波动与长期趋势。

(3) 交互缺失问题：基本面特征一般滞后报告期 1 个月发布，即存在一定的滞后效应，但传统数据处理方法却忽略这一滞后效应。

根据上述常规量价特征与基本面特征处理存在的问题，本文设计了多尺度信息融合框架。其中多尺度信息融合框架包含**微观尺度**、**中观尺度**以及**宏观尺度**。其中微观尺度主要处理原始日频量价数据，通过时间卷积捕捉短期波动模式；中观尺度主要对日频量价数据做中期时间长度（如周度）聚合处理，分析量价信息的中期趋势；宏观尺度主要处理基本面特征，保持原有基本面特征频次。

图 19：多尺度信息融合自适应：微、中、宏观尺度做基本面及量价信息处理



数据来源：西南证券整理

在原有的微观尺度、中观尺度以及宏观尺度信息处理后，本文引入**跨尺度注意力机制**实现不同尺度信息间的“对话”，并进行跨尺度注意力门控融合，并于原始量价特征进行残差融合，保证梯度下降传播。具体操作步骤如下：

- 1) 以微观尺度特征作为查询 (Query)，中观、宏观尺度特征作为键值对 (Key-Value)。
- 2) **计算注意力分数**：利用查询 (Query) 分别与键 (Key) 计算注意力分数，并进行堆叠；
- 3) **计算尺度门控权重**：拼接三个尺度的特征并通过门控网络生成三个尺度的门控权重，动态学习并决定在融合时，每个尺度的重要性占比；
- 4) **应用门控权重，融合注意力分数，计算注意力权重**：将第 3 步得到的门控权重应用到第 2 步得到的三个注意力分数上，融合成一个加权的注意力分数矩阵，将融合后的注意力分数转换为概率分布 (softmax)；
- 5) **上下文聚合**：合并三个尺度的值 (Value)，使用门控权重对合并值加权，生成融合后的值向量，并于门控权重相乘，可理解为用注意力权重对所有时间步的值向量进行加权平均，为每一个时间步 T 生成一个包含多尺度信息的上下文向量；
- 6) **残差融合**：组合多头并输出投影至隐藏层维度，并与原始量价特征进行残差融合。

除第一章本文介绍使用的 14 个量价特征以外，本文选择如下 8 个特征作为模型训练使用的基本面特征：

表 14：模型基本面特征列表

序号	特征中文名称	特征助记符	特征计算方法
1	EPS 同比	EPSYOY	EPS 同比增长率
2	归母净利润同比	NIYOY	归母净利润同比增长率
3	ROE 环比	ROE_QoQ	ROE 季度环比增长率
4	市净率	PB	市净率
5	市盈率	PE	市盈率
6	市销率	PS	市销率
7	股息率	Dividend	股息率
8	归母净利润/负债	NPL	归母净利润与负债的比值

数据来源：西南证券整理

上述多尺度信息融合模块在整体的 Transformer 模型中整体位于特征嵌入层后介入，Transformer 层前的位置，负责输入特征信息处理，以实现 Transformer 自适应模型中的信息融合自适应功能。

4.1 信息融合自适应 Transformer 模型测试

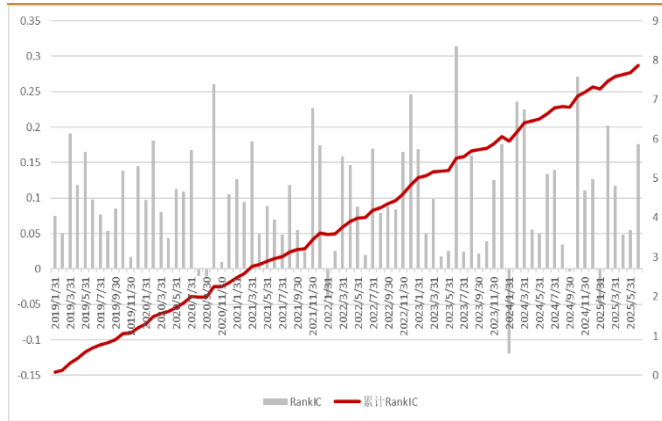
本节中测试在基础 Transformer 模型中仅加入多尺度信息融合自适应的 Transformer 模型的选股因子效果 (MF_Transformer, Multi-scale Information Fusion Transformer)。

自 2019 年 1 月至 2025 年 7 月，全 A 范围内 MF 模型因子月均 IC 为 10.09%，多头组合年化收益率为 30.22%，月均单边换手率 0.82X。整体来看，引入基本面特征并采取多尺度信息融合机制下的 Transformer 模型因子 IC 及多头组合表现略优于基础 Transformer 模型，且因子换手率有所降低。

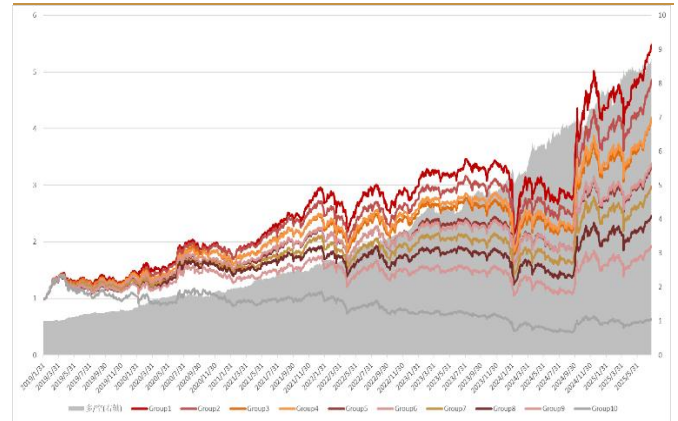
表 15: MF_Transformer 因子回测结果

	IC	ICIR	Top10%组合年化收益率	年化波动	信息比率	最大回撤率	单边月均换手率
MF_Transformer	0.1009***	4.40	30.22%	24.80%	1.22	37.90%	0.82X

数据来源: wind、恒生聚源、西南证券整理, 截至 2025 年 7 月 31 日

图 20: MF_Transformer 因子 IC


数据来源: wind、西南证券整理, 截至 2025 年 7 月 31 日

图 21: MF_Transformer 因子分组-绝对净值


数据来源: wind、西南证券整理, 截至 2025 年 7 月 31 日

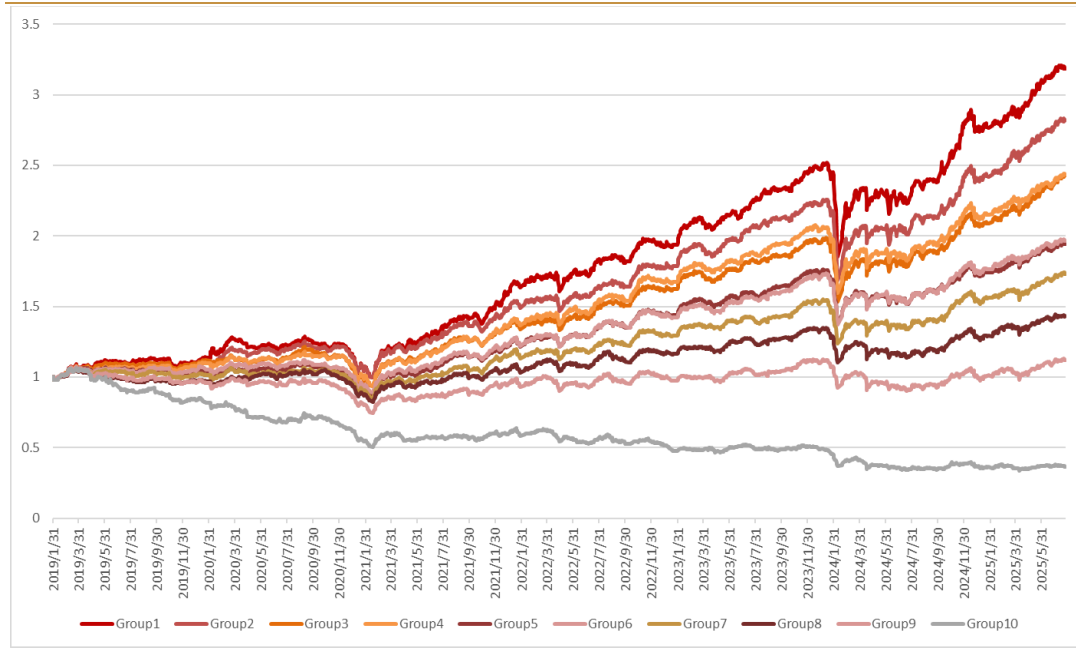
从分组角度来看，多头组合（组 1）年化收益率 30.22%，最大回撤 37.90%；空头组合（组 10）年化收益率-6.91%，最大回撤 71.65%；多空组合（组 1 / 组 10）年化收益率 39.89%，最大回撤 9.24%，MF_Transformer 模型具备一定单调性。纵向对比来看，MF_Transformer 相对于基础 Transformer 模型在空头与多头均有提升；横向对比来看，MF_Transformer 相对于 SA_Transformer 以及 DPE_Transformer 模型在空头与多头同样具备一定的提升效果。

表 16: MF_Transformer 因子分组表现

	年化收益率	年化波动	信息比率	最大回撤率
Group1	30.22%	24.80%	1.22	37.90%
Group2	27.79%	24.45%	1.14	38.12%
Group3	24.87%	24.21%	1.03	36.94%
Group4	24.89%	24.43%	1.02	36.55%
Group5	20.61%	24.46%	0.84	35.63%
Group6	20.83%	24.55%	0.85	34.64%
Group7	18.46%	24.79%	0.74	35.01%
Group8	15.02%	25.18%	0.60	35.56%
Group9	10.71%	25.86%	0.41	39.94%
Group10	-6.91%	27.57%	-0.25	71.65%
多/空	39.89%	12.39%	3.22	9.24%

数据来源: wind、恒生聚源、西南证券整理, 截至 2025 年 7 月 31 日

从因子分组的相对净值表现来看，相对万得全 A，MF_Transformer 模型因子在 2024 年 1 月底同样出现了较大的超额回撤，但随后回撤快速修复，且在 2024 年 9 月底的上涨行情（0924 行情）内并未出现较为明显的超额回撤。

图 22: MF_Transformer 因子分组测试-相对净值


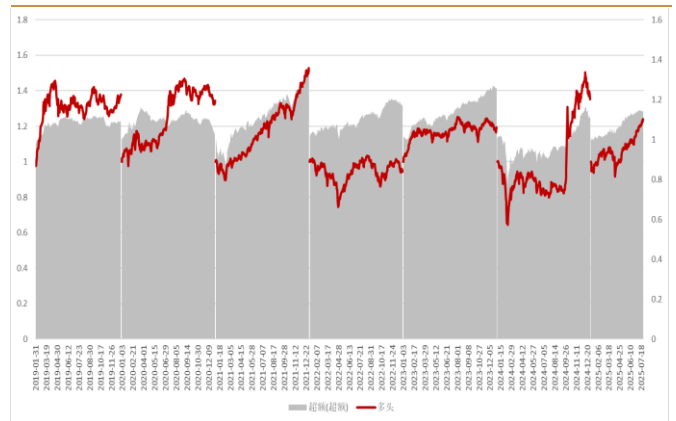
数据来源: wind、西南证券整理, 截至 2025 年 7 月 31 日

分年度来看, 多头组合自 2019 年至 2025 年 7 月, 7 年均跑赢万得全 A 指数, 其中 2021 年相对万得全 A 年化超额 32.24%, 超额最大回撤 11.98%; 2023 全年相对万得全 A 年化超额收益率 24.15%, 超额最大回撤仅 4.35%。此外, 因子今年以来超额表现较好, 2025 年初至 7 月 31 日, 累计超额收益率 13.90%, 超额最大回撤 2.44%。相较于基础 Transformer, MF_Transformer 模型在 2024 年表现有所提升, 其多头年化超额收益率为 11.08%, 超额回撤 25.64%。但整体来看, MF_Transformer 在 2020 年的表现较差, 相较于万得全 A 超额仅 2.35%。

表 17: MF_Transformer 因子多头分年度超额表现

年份	年化收益率	年化波动率	信息比率	最大回撤率
2019	8.78%	8.31%	1.06	5.34%
2020	2.35%	9.27%	0.25	11.80%
2021	35.52%	15.01%	2.37	12.89%
2022	16.97%	10.32%	1.64	7.52%
2023	25.60%	6.73%	3.80	4.43%
2024	11.08%	20.33%	0.54	25.64%
202507	13.90%	8.62%	1.61	2.44%

数据来源: wind、ifind、西南证券整理, 截至 2025 年 7 月 31 日

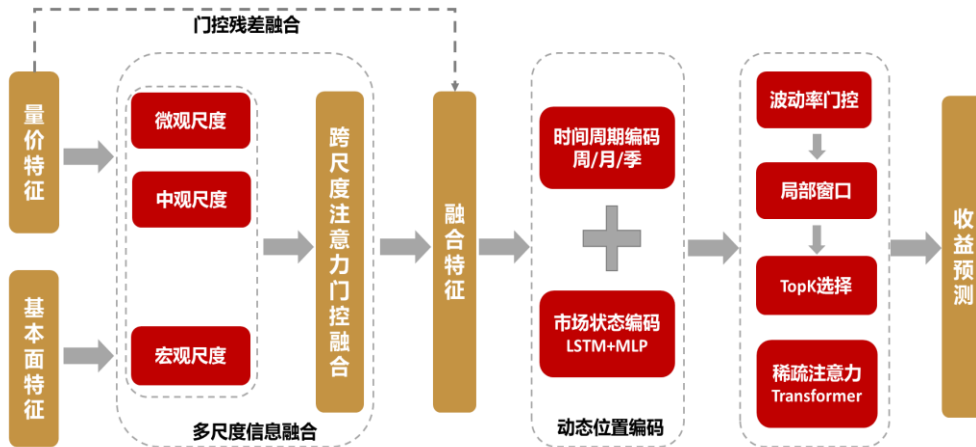
图 23: MF_Transformer 因子多头分年度净值


数据来源: wind、ifind、西南证券整理, 截至 2025 年 7 月 31 日

5 模型消融测试

基于前文提到的三个自适应优化模块，本文在 Transformer 模型的基础上构建了 DAFAT (Dynamic Adaptive Fusion Attention Transformer) 模型，并分别测试上述三个不同的自适应优化模块（动态位置编码 DPE、稀疏注意力 SA、多频信息融合 MF）在任意组合情况下相对于传统 Transformer 模型的效果优化情况。

图 24：DAFAT 模型结构示意图



数据来源：wind、ifind、西南证券整理，截至 2025 年 7 月 31 日

在上述三个自适应优化模块两两组合的情况下，模型表现相较于传统 Transformer 模型具备一定的削弱效果。其中 MF_DPE、SA_DPE 以及 SA_MF 相较于基础 Transformer 模型在多头组合方面和 IC 方面并没有特别显著的提升，其中该三者模型因子 IC 分别为 9.77%、8.74%以及 9.35%，多头组合年化收益率分别为 26.44%、26.97%以及 28.26%。

表 18：自适应优化模块模型消融测试

模型名称	IC	ICIR	TOP10% 年化收益率	年化波动	信息比率	TOP10% 最大回撤率	换手率
Transformer	0.0942***	4.03	28.70%	24.81%	1.16	39.43%	0.87X
DPE	0.0954***	4.44	30.60%	24.97%	1.23	39.41%	0.86X
SA	0.0998***	4.24	29.64%	24.74%	1.20	38.62%	0.87X
MF	0.1009***	4.40	30.22%	24.80%	1.22	37.90%	0.82X
MF_DPE	0.0977***	4.18	26.44%	25.00%	1.06	39.45%	0.81X
SA_DPE	0.0874***	4.30	26.97%	24.99%	1.08	38.67%	0.86X
SA_MF	0.0935***	4.13	28.26%	24.93%	1.13	38.52%	0.82X
DAFAT	0.1107***	5.21	32.30%	24.11%	1.30	33.00%	0.87X

数据来源：wind、恒生聚源、ifind、西南证券整理，截至 2025 年 7 月 31 日

从各模型相对万得全 A 的相对净值表现来看，集成三类自适应优化模块的 DAFAT 模型在近几年的效果有明显提升，自 2024 年 1 月底量化模型大幅回撤后，DAFAT 模型超额弹性显著增强，且 2024 年 1 月底期间回撤相对于其余几个模型更低。2024 年以前，DPE 模型相对于其他模型超额收益表现更佳。

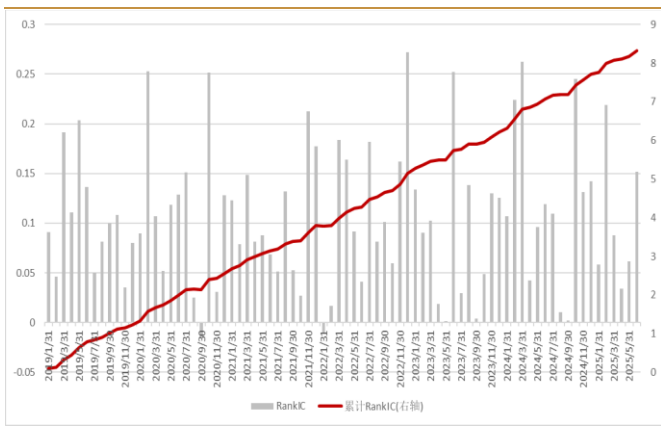
图 25: Transformer 自适应模块消融测试-相对净值



数据来源: wind、ifind、西南证券整理, 截至 2025 年 7 月 31 日

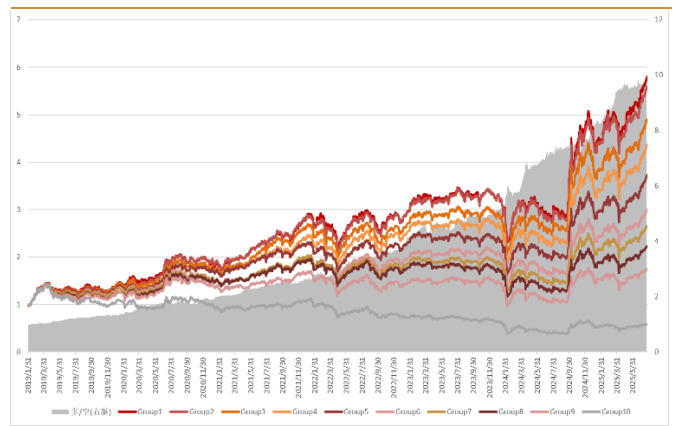
此外, 同时融合了三个自适应优化模块的 DAFAT 模型相较于基础 Transformer 以及其他不完全自适应优化的 Transformer 模型均呈现出较好的优化效果。其中 DAFAT 自 2019 年 1 月至 2025 年 7 月因子 IC 均值为 11.07%, 多头组合年化收益率 32.30%, 最大回撤率 33.00%。

图 26: DAFAT 因子 IC



数据来源: wind、ifind、西南证券整理, 截至 2025 年 7 月 31 日

图 27: DAFAT 因子分组-绝对净值



数据来源: wind、ifind、西南证券整理, 截至 2025 年 7 月 31 日

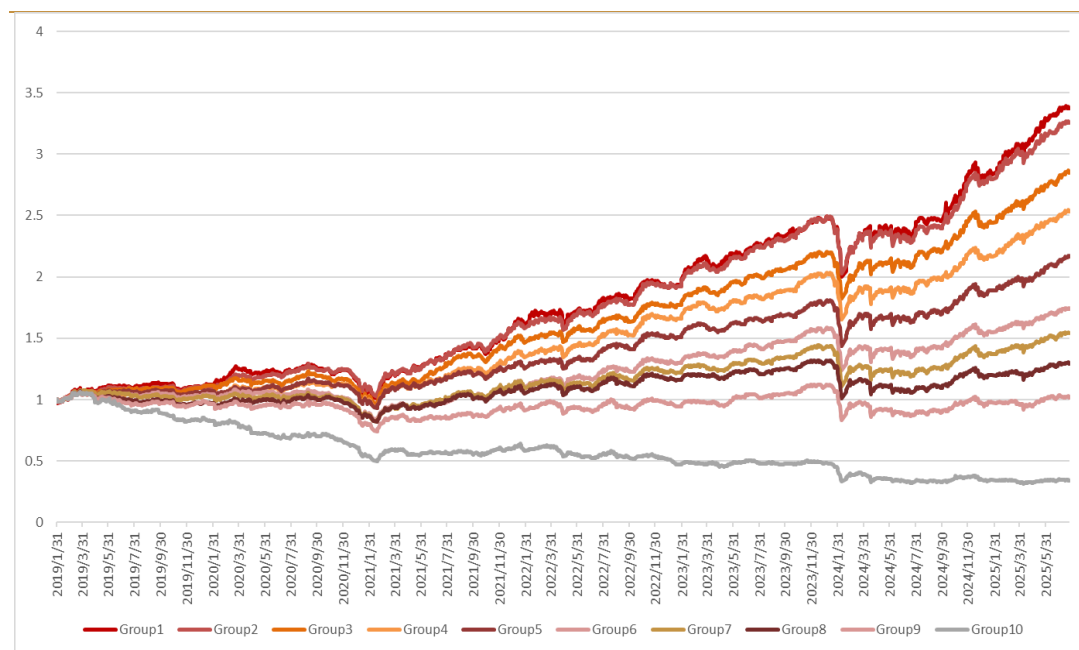
从分组角度来看, 多头组合 (组 1) 年化收益率 32.30%, 最大回撤 33.00%; 空头组合 (组 10) 年化收益率 -7.96%, 最大回撤 73.13%; 多空组合 (组 1 / 组 10) 年化收益率 42.66%, 最大回撤 9.07%, DAFAT 模型具备一定单调性, 且多头组合与空头组合均优于基础 Transformer 模型。但整体来看, DAFAT 因子组一、组二分化程度并不明显。

表 19：DAFAT 因子分组表现

	年化收益率	年化波动	信息比率	最大回撤率
Group1	32.30%	24.11%	1.30	33.00%
Group2	30.65%	23.91%	1.28	32.42%
Group3	28.01%	23.86%	1.17	32.73%
Group4	25.68%	24.12%	1.06	33.67%
Group5	22.67%	24.49%	0.93	35.19%
Group6	18.59%	24.77%	0.75	35.40%
Group7	16.36%	25.05%	0.65	37.88%
Group8	13.25%	25.51%	0.52	41.24%
Group9	9.13%	26.18%	0.35	43.64%
Group10	-7.96%	27.92%	-0.29	73.13%
多/空	42.66%	11.91%	3.58	9.07%

数据来源：wind、恒生聚源、ifind、西南证券整理，截至2025年7月31日

从因子分组的相对净值表现来看，相对万得全 A，MF_Transformer 模型因子在 2024 年 1 月底同样出现了较大的超额回撤，但随后回撤快速修复，且在 2024 年 9 月底的上涨行情（0924 行情）内并未出现较为明显的超额回撤。

图 28：DAFAT 因子分组测试-相对净值


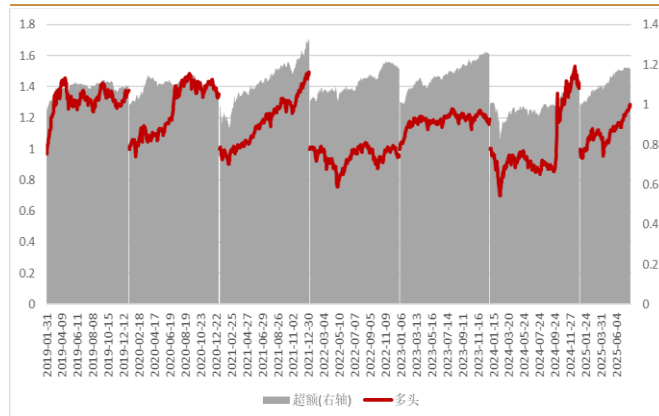
数据来源：wind、ifind、西南证券整理，截至2025年7月31日

分年度来看，多头组合自 2019 年至 2025 年 7 月，7 年均跑赢万得全 A 指数，其中 2021 年相对万得全 A 年化超额 32.87%，超额最大回撤 12.14%；2023 全年相对万得全 A 年化超额收益率 25.41%，超额最大回撤仅 5.09%。此外，因子 2025 年初至 7 月 31 日累计超额收益率 18.04%，超额最大回撤 2.35%。但同样 DAFAT 在 2019 与 2020 年超额收益表现较差，相对万得全 A 年化超额收益率仅 8.81% 与 3.23%。

表 20: DAFAT 因子多头分年度超额表现

年份	年化收益率	年化波动率	信息比率	最大回撤率
2019	8.81%	8.41%	1.05	5.29%
2020	3.23%	8.98%	0.36	11.20%
2021	32.87%	14.54%	2.26	12.14%
2022	18.08%	10.21%	1.77	7.21%
2023	25.41%	7.11%	3.57	5.09%
2024	13.53%	14.84%	0.91	19.05%
202507	18.04%	8.31%	2.17	2.35%

数据来源: wind、ifind、西南证券整理, 截至 2025 年 7 月 31 日

图 29: DAFAT 因子多头分年度净值


数据来源: wind、ifind、西南证券整理, 截至 2025 年 7 月 31 日

6 指数增强策略

6.1 沪深 300 指数增强策略

在本节中, 本文基于 DAFAT 模型构建沪深 300 指数增强策略。首先在构建指数增强策略之前, 本文分别测试了前文中 DAFAT 消融测试 8 个模型分别在沪深 300 指数成分股内表现。

整体来看, 在单个自适应优化模块下, 多尺度信息融合自适应 (MF_transformer) 优化下相对基础 Transformer 模型表现相对最好; 在两个自适应模块优化的情况下, MF_DPE、SA_DPE 以及 SA_MF 相对基础 Transformer 模型并没有较大的提升。但在三个自适应模块的优化下, DAFAT 模型相对于其余模型在沪深 300 成分股内均有一定程度的提升。自 2019 年 1 月至 2025 年 7 月, DAFAT 在沪深 300 成分股内 IC 均值为 6.24%, 相对沪深 300 指数年化超额收益率 17.17%。

表 21: 沪深 300 指数成分股因子测试

模型名称	IC	ICIR	年化收益率	年化波动	信息比率	最大回撤率	年化超额	换手率
transformer	0.0548***	2.00	21.65%	20.75%	1.04	25.29%	14.78%	0.89X
DPE_transformer	0.0539***	1.75	19.37%	21.18%	0.91	28.42%	12.75%	0.89X
SA_transformer	0.0484***	1.77	19.03%	20.63%	0.92	25.37%	12.27%	0.89X
MF_transformer	0.0560***	1.85	16.74%	20.16%	0.83	23.60%	10.04%	0.85X
MF_DPE	0.0525***	1.74	19.02%	20.34%	0.94	22.81%	12.28%	0.84X
SA_DPE	0.0458***	1.59	18.31%	21.23%	0.86	30.88%	11.77%	0.88X
SA_MF	0.0469***	1.59	15.49%	20.16%	0.77	26.22%	8.86%	0.85X
DAFAT	0.0624***	2.33	23.99%	20.61%	0.97	25.26%	17.17%	0.89X

数据来源: wind、恒生聚源、ifind、西南证券整理, 截至 2025 年 7 月 31 日

其中，本文沪深 300 指数增强策略回测设定如下：

- 1) 个股权重偏离：个股权重偏离不超过 1%；
- 2) 行业权重偏离：申万一级行业偏离不超过 5%；
- 3) 成分股权重限制：沪深 300 成分股权重不少于 80%；
- 4) 交易费率：单边千分之二；
- 5) 回测时间 2019 年 1 月至今，月频调仓，调仓当日以 wwap 价格成交。

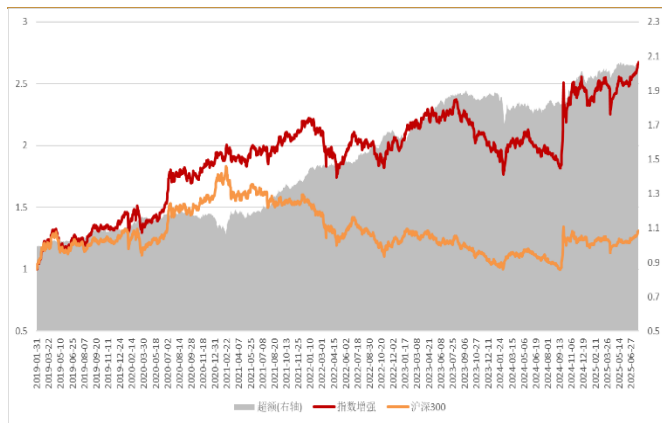
表 22：沪深 300 指数增强策略回测结果

	年化收益率	年化波动率	信息比率	最大回撤率	年化超额收益率
DAFAT	17.07%	19.13%	0.89	-25.43%	12.65%

数据来源：wind、恒生聚源、ifind、西南证券整理，截至 2025 年 7 月 31 日

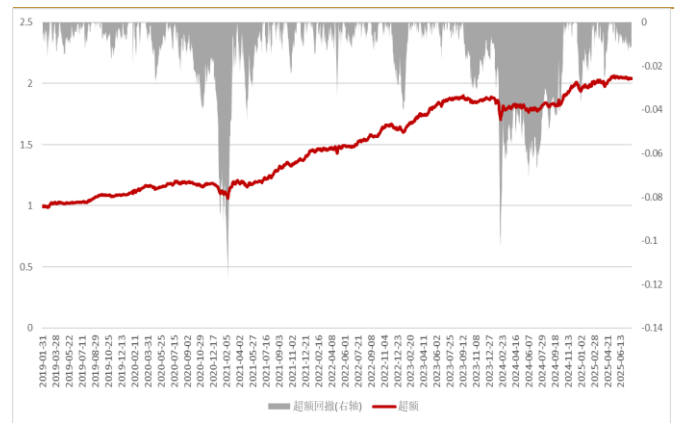
整体来看，DAFAT 模型沪深 300 指数增强策略年化超额收益率 12.65%。根据 DAFAT 指数增强超额表现来看，历史上 DAFAT 沪深 300 指数增强在 2021 年 2 月及 2024 年 2 月至 7 月出现了较大幅度的超额回撤。

图 30：DAFAT 沪深 300 指数增强表现



数据来源：wind、ifind、西南证券整理，截至 2025 年 7 月 31 日

图 31：DAFAT 沪深 300 指数增强超额滚动回撤



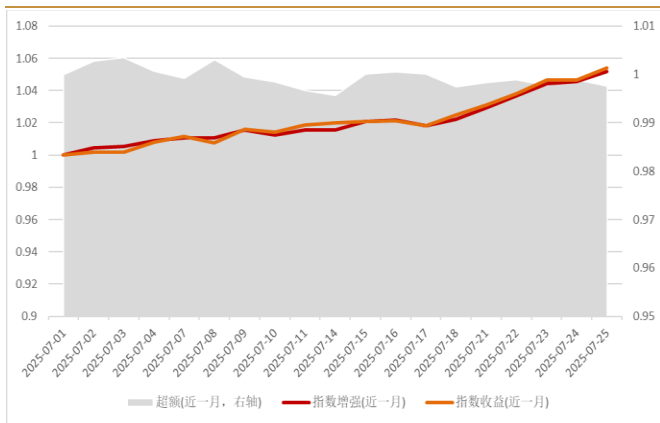
数据来源：wind、ifind、西南证券整理，截至 2025 年 7 月 31 日

分区间来看，截至 2025 年 7 月 31 日，DAFAT 模型沪深 300 指数增强近一个月、三个月、一年、三年超额收益率分别为 -0.25%、0.20%、14.76% 以及 10.39%（超过一年计算年化收益率，不足一年计算累计收益率）。

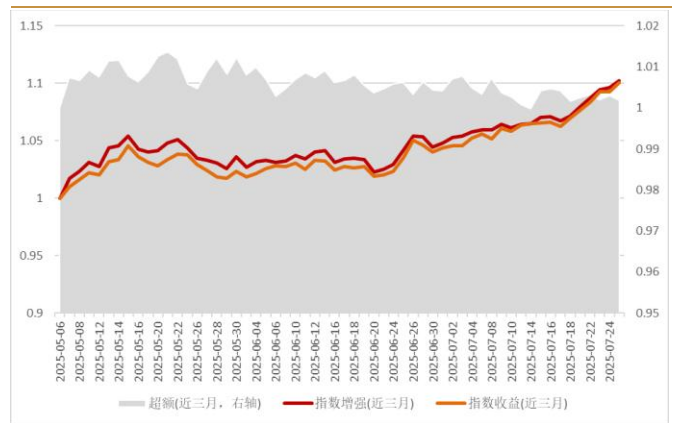
表 23：DAFAT 沪深 300 指数增强策略分区间回测结果

	年化收益率	年化波动率	信息比率	最大回撤率	年化超额收益率
指数增强(近一年)	37.90%	24.15%	1.57	-12.80%	14.76%
指数增强(近三年)	10.21%	18.79%	0.54	-25.43%	10.39%
指数增强(近一月)	5.16%	5.11%	1.01	-0.35%	-0.25%
指数增强(近三月)	10.24%	9.66%	1.06	-2.93%	0.20%

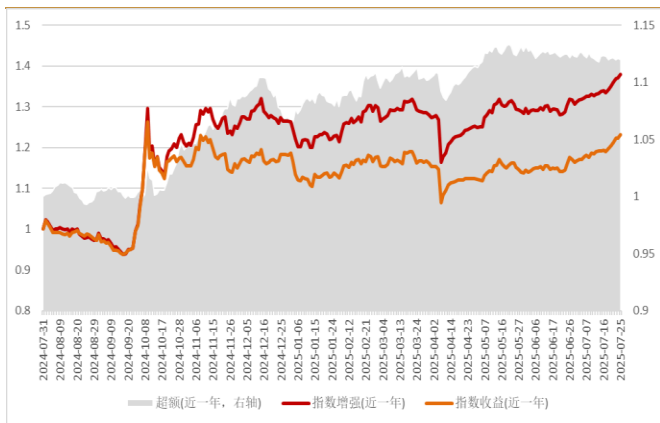
数据来源：wind、恒生聚源、ifind、西南证券整理，截至 2025 年 7 月 31 日

图 32: DAFAT 沪深 300 指数增强表现 (近一月)


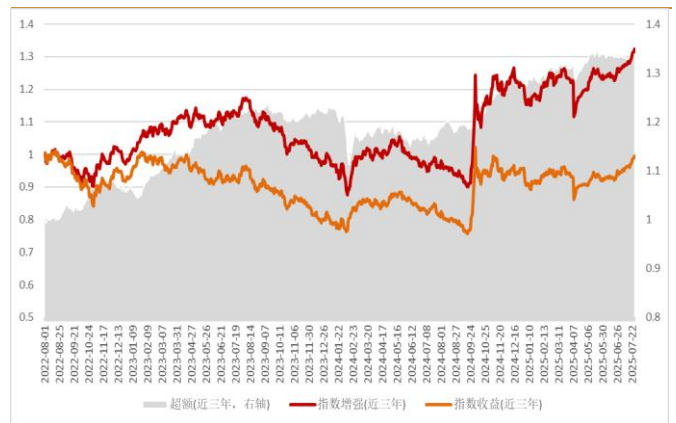
数据来源: wind、ifind、西南证券整理, 截至 2025 年 7 月 31 日

图 33: DAFAT 沪深 300 指数增强表现 (近三月)


数据来源: wind、ifind、西南证券整理, 截至 2025 年 7 月 31 日

图 34: DAFAT 沪深 300 指数增强表现 (近一年)


数据来源: wind、ifind、西南证券整理, 截至 2025 年 7 月 31 日

图 35: DAFAT 沪深 300 指数增强表现 (近三年)


数据来源: wind、ifind、西南证券整理, 截至 2025 年 7 月 31 日

6.2 中证 1000 指数增强策略

在本节中, 本文基于 DAFAT 模型构建中证 1000 指数增强策略。首先在构建指数增强策略之前, 本文分别测试了前文中 DAFAT 类消融测试的 8 个模型分别在中证 1000 指数成分股内的表现。

横向对比来看, 在单个自适应模块优化的情况下, 动态位置编码自适应模块为 Transformer 模型带来了较大程度的提升。自 2019 年 1 月至 2025 年 7 月, DPE_Transformer 模型在中证 1000 成分股内 IC 均值为 9.22%, 相对中证 1000 指数年化超额收益率 11.61%, 优于 SA_Transformer、MF_Transformer 模型。此外在双重自适应模块优化的情况下, 模型表现整体并没有给基础 Transformer 模型带来明显提升; 最后在三重自适应模块优化的情况下, DAFAT 因子 IC 均值为 9.52%, 多头组合相对中证 1000 年化超额收益率 13.42%, 表现相对其余模型有一定程度提升。

表 24：中证 1000 指数成分股因子测试

	IC	ICIR	TOP10% 年化收益率	年化波动率	信息比率	最大回撤率	TOP10% 年化超额	换手率
transformer	0.0827***	3.69	21.96%	25.00%	0.88	37.22%	9.96%	0.87X
DPE_transformer	0.0922***	4.36	24.10%	25.32%	1.11	36.87%	11.61%	0.87X
SA_transformer	0.0821***	3.74	22.65%	25.03%	0.90	37.58%	10.57%	0.87X
MF_transformer	0.0859***	3.68	22.54%	24.75%	0.91	37.43%	10.35%	0.81X
MF_DPE	0.0848***	3.67	20.95%	25.26%	0.83	36.57%	9.10%	0.81X
SA_DPE	0.0819***	3.86	21.94%	25.49%	0.86	38.35%	10.11%	0.87X
SA_MF	0.0791***	3.37	22.20%	25.02%	0.89	36.78%	10.13%	0.82X
DAFAT	0.0952***	4.58	25.80%	24.67%	1.01	34.03%	13.42%	0.87X

数据来源：wind、恒生聚源、ifind、西南证券整理，截至 2025 年 7 月 31 日

其中，本文中证 1000 指数增强策略回测设定如下：

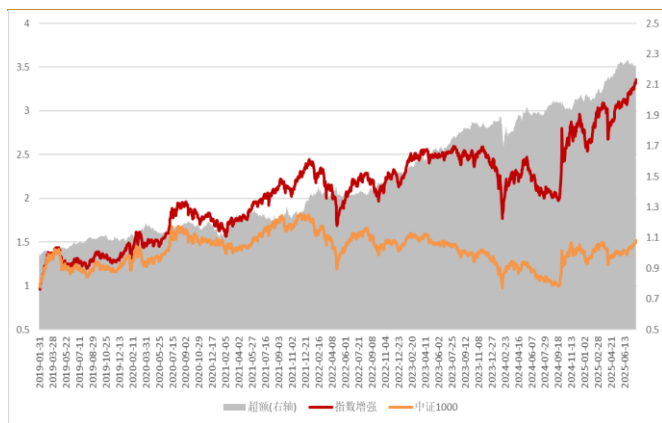
- 1) 个股权重偏离：个股权重偏离不超过 1%；
- 2) 行业权重偏离：申万一级行业偏离不超过 5%；
- 3) 成分股权重限制：中证 1000 成分股权重不少于 80%；
- 4) 交易费率：单边千分之二；
- 5) 回测时间 2019 年 1 月至今，月频调仓，调仓当日以 wap 价格成交。

表 25：中证 1000 指数增强策略回测结果

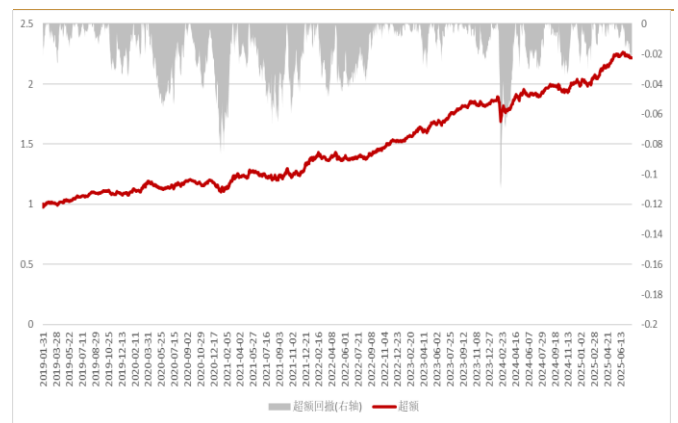
	年化收益率	年化波动率	信息比率	最大回撤	年化超额收益率
DAFAT	21.46%	24.21%	0.89	-31.82%	14.57%

数据来源：wind、恒生聚源、ifind、西南证券整理，截至 2025 年 7 月 31 日

整体来看，DAFAT 模型中证 1000 指数增强策略年化超额收益率 14.57%。根据指数增强超额收益滚动回撤表现来看，DAFAT 模型中证 1000 指数增强超额与 2021 年 2 月及 2024 年 1 月末出现较大幅度的超额回撤。

图 36：DAFAT 中证 1000 指数增强表现


数据来源：wind、ifind、西南证券整理，截至 2025 年 7 月 31 日

图 37：DAFAT 中证 1000 指数增强相对净值


数据来源：wind、ifind、西南证券整理，截至 2025 年 7 月 31 日

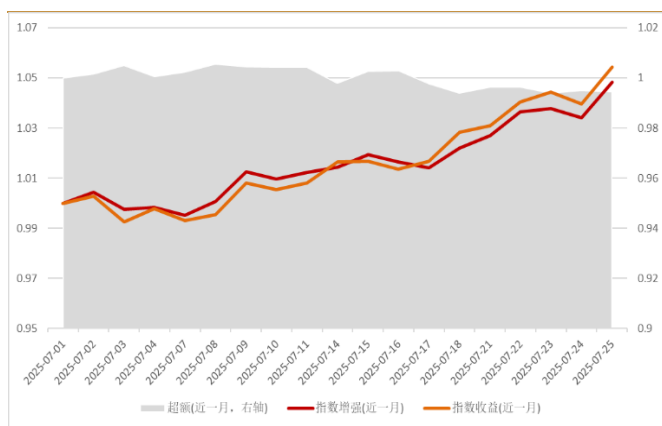
分区间来看，截至 2025 年 7 月 31 日，DAFAT 模型中证 1000 指数增强近一个月、三个月、一年、三年超额收益率分别为 -0.60%、2.41%、20.85% 以及 16.79%（超过一年计算年化收益率，不足一年计算累计收益率）。

表 26：DAFAT 中证 1000 指数增强策略分区间回测结果

	年化收益率	年化波动率	信息比率	最大回撤	年化超额收益率
指数增强(近一年)	63.35%	30.51%	2.08	-14.32%	20.85%
指数增强(近三年)	14.71%	24.67%	0.60	-31.82%	16.79%
指数增强(近一月)	4.83%	8.95%	0.54	-0.89%	-0.60%
指数增强(近三月)	15.04%	12.94%	1.16	-2.47%	2.41%

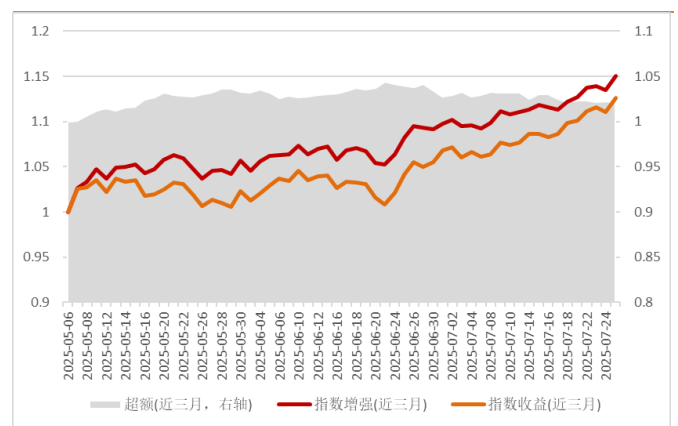
数据来源：wind、恒生聚源、ifind、西南证券整理，截至 2025 年 7 月 31 日

图 38：DAFAT 中证 1000 指数增强表现（近一月）



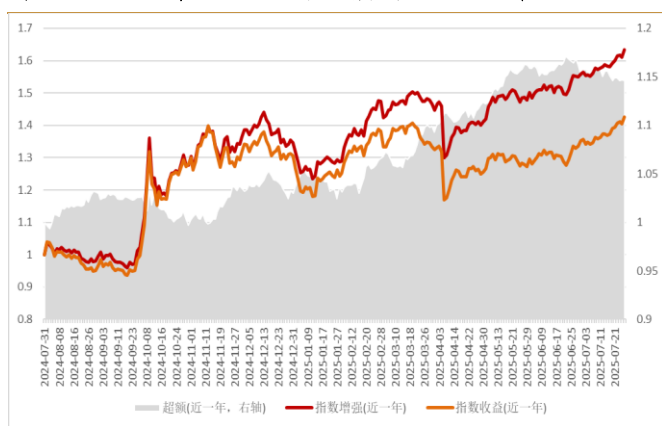
数据来源：wind、ifind、西南证券整理，截至 2025 年 7 月 31 日

图 39：DAFAT 中证 1000 指数增强表现（近三月）



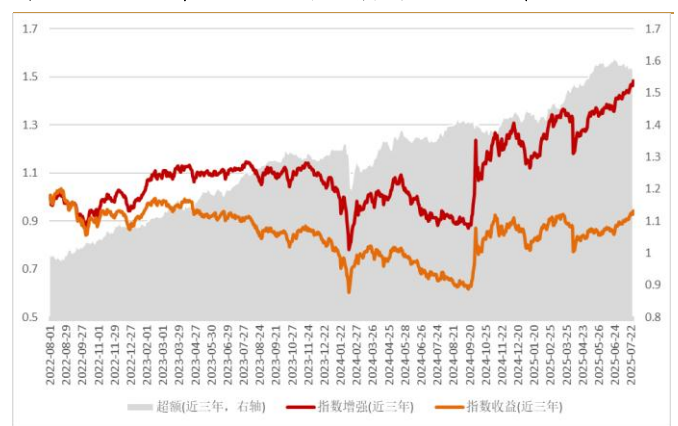
数据来源：wind、ifind、西南证券整理，截至 2025 年 7 月 31 日

图 40：DAFAT 中证 1000 指数增强表现（近一年）



数据来源：wind、ifind、西南证券整理，截至 2025 年 7 月 31 日

图 41：DAFAT 中证 1000 指数增强表现（近三年）



数据来源：wind、ifind、西南证券整理，截至 2025 年 7 月 31 日

7 总结与展望

DAFAT（动态自适应融合注意力 Transformer）模型通过三大核心创新实现了金融选股领域的突破性进展。在**动态位置编码**方面，模型创新性地融合了时间周期、市场状态的双重编码机制，通过门控网络动态调整权重分配，改善了传统位置编码在金融时序中的周期失配、状态盲区问题。在稀疏注意力机制上，DAFAT 模型尝试引入波动率门控、局部注意力窗口和 Top-k 选择三重稀疏化策略，致力于改善传统注意力的计算效率瓶颈和噪声敏感问题。波动率阈值过滤，聚焦关键市场波动时段，短期局部窗口捕捉金融短期依赖特性，30%的 Top-k 连接保留显著提升了计算效率。**多尺度信息融合模块**则通过微观（日频）、中观（周频）、宏观（月频）三级处理架构，解决了不同频率特征融合的核心难题。跨尺度注意力机制创新性地以量价特征为 Query，基本面特征为 Key/Value，配合门控残差融合。上述三个自适应优化模块的加入后，模型相对于 Transformer 模型均有不同程度的提升，且三个自适应模块优化的情况下，DAFAT 模型相较于基础 Transformer 以及其他不完全自适应优化的 Transformer 模型均呈现出较好的优化效果。其中 DAFAT 自 2019 年 1 月至 2025 年 7 月因子 IC 均值为 11.07%，多头组合年化收益率 32.30%，最大回撤率 33.00%。

该报告提到的 DAFAT 模型为基于 Transformer 模型，分别从位置编码、注意力机制以及信息融合机制三个方向进行的模型端改进。展望未来，我们会进一步尝试从以下几个维度进一步突破：首先，在模型架构层面，可探索时空图神经网络（STGNN）的引入，将行业关联、供应链关系等图结构数据纳入建模框架，解决传统 Transformer 在截面关系建模的不足。同时，尝试研发条件计算机制，根据市场波动状态动态调整模型复杂度，在平稳期启用轻量化子网络以提升推理效率。其次，在数据融合维度，重点突破另类数据整合技术，开发新闻情感等另类数据的多模态融合模块，通过跨模态注意力对齐文本与数值特征；建立宏观-微观数据桥接网络，实现货币政策、行业政策等宏观信号到个股影响的量化传导。

8 风险提示

报告对应的相关结论完全基于公开的历史数据进行算法构建、统计以及计算，文中部分数据有一定滞后性，同时也存在第三方数据提供不准确或者缺失等风险；策略效果结论仅针对于回测区间得出，并不预示其未来表现，也不能保证未来的可持续性，亦不构成投资收益的保证或投资建议。

分析师承诺

本报告署名分析师具有中国证券业协会授予的证券投资咨询执业资格并注册为证券分析师，报告所采用的数据均来自合法合规渠道，分析逻辑基于分析师的职业理解，通过合理判断得出结论，独立、客观地出具本报告。分析师承诺不曾因，不因，也将不会因本报告中的具体推荐意见或观点而直接或间接获取任何形式的补偿。

投资评级说明

报告中投资建议所涉及的评级分为公司评级和行业评级（另有说明的除外）。评级标准为报告发布日后 6 个月内的相对市场表现，即：以报告发布日后 6 个月内公司股价（或行业指数）相对同期相关证券市场代表性指数的涨跌幅作为基准。其中：A 股市场以沪深 300 指数为基准，新三板市场以三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的）为基准；香港市场以恒生指数为基准；美国市场以纳斯达克综合指数或标普 500 指数为基准。

公司评级	买入：未来 6 个月内，个股相对同期相关证券市场代表性指数涨幅在 20% 以上
	持有：未来 6 个月内，个股相对同期相关证券市场代表性指数涨幅介于 10% 与 20% 之间
	中性：未来 6 个月内，个股相对同期相关证券市场代表性指数涨幅介于 -10% 与 10% 之间
	回避：未来 6 个月内，个股相对同期相关证券市场代表性指数涨幅介于 -20% 与 -10% 之间
	卖出：未来 6 个月内，个股相对同期相关证券市场代表性指数涨幅在 -20% 以下
行业评级	强于大市：未来 6 个月内，行业整体回报高于同期相关证券市场代表性指数 5% 以上
	跟随大市：未来 6 个月内，行业整体回报介于同期相关证券市场代表性指数 -5% 与 5% 之间
	弱于大市：未来 6 个月内，行业整体回报低于同期相关证券市场代表性指数 -5% 以下

重要声明

西南证券股份有限公司（以下简称“本公司”）具有中国证券监督管理委员会核准的证券投资咨询业务资格。

本公司与作者在自身所知知情范围内，与本报告中所评价或推荐的证券不存在法律法规要求披露或采取限制、静默措施的利益冲突。

《证券期货投资者适当性管理办法》于 2017 年 7 月 1 日起正式实施，本报告仅供本公司签约客户使用，若您并非本公司签约客户，为控制投资风险，请取消接收、订阅或使用本报告中的任何信息。本公司也不会因接收人收到、阅读或关注自媒体推送本报告中的内容而视其为客户。本公司或关联机构可能会持有报告中提到的公司所发行的证券并进行交易，还可能为这些公司提供或争取提供投资银行或财务顾问服务。

本报告中的信息均来源于公开资料，本公司对这些信息的准确性、完整性或可靠性不作任何保证。本报告所载的资料、意见及推测仅反映本公司于发布本报告当日的判断，本报告所指的证券或投资标的的价格、价值及投资收入可升可跌，过往表现不应作为日后的表现依据。在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告，本公司不保证本报告所含信息保持在最新状态。同时，本公司对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。

本报告仅供参考之用，不构成出售或购买证券或其他投资标的的要约或邀请。在任何情况下，本报告中的信息和意见均不构成对任何个人的投资建议。投资者应结合自己的投资目标和财务状况自行判断是否采用本报告所载内容和信息并自行承担风险，本公司及雇员对投资者使用本报告及其内容而造成的一切后果不承担任何法律责任。

须注明出处为“西南证券”，且不得对本报告及附录进行有悖原意的引用、删节和修改。未经授权刊载或者转发本报告及附录的，本公司将保留向其追究法律责任的权利。

请务必阅读正文后的重要声明部分

西南证券研究院

上海

地址：上海市浦东新区陆家嘴 21 世纪大厦 10 楼

邮编：200120

北京

地址：北京市西城区金融大街 35 号国际企业大厦 A 座 8 楼

邮编：100033

深圳

地址：深圳市福田区益田路 6001 号太平金融大厦 22 楼

邮编：518038

重庆

地址：重庆市江北区金沙门路 32 号西南证券总部大楼 21 楼

邮编：400025

西南证券机构销售团队

区域	姓名	职务	座机	手机	邮箱
上海	蒋诗烽	院长助理、研究销售部经理、 上海销售主管	021-68415309	18621310081	jsf@swsc.com.cn
	崔露文	销售岗	15642960315	15642960315	clw@swsc.com.cn
	李煜	销售岗	18801732511	18801732511	yfliyu@swsc.com.cn
	汪艺	销售岗	13127920536	13127920536	wywf@swsc.com.cn
	戴剑箫	销售岗	13524484975	13524484975	daijx@swsc.com.cn
	李嘉隆	销售岗	15800507223	15800507223	ljlong@swsc.com.cn
	欧若诗	销售岗	18223769969	18223769969	ors@swsc.com.cn
	蒋宇洁	销售岗	15905851569	15905851569	jjj@swsc.com.c
	贾文婷	销售岗	13621609568	13621609568	jiawent@swsc.com.cn
	张嘉诚	销售岗	18656199319	18656199319	zhangjc@swsc.com.cn
毛玮琳	销售岗	18721786793	18721786793	mwl@swsc.com.cn	
北京	李杨	北京销售主管	18601139362	18601139362	yfly@swsc.com.cn
	张岚	销售岗	18601241803	18601241803	zhanglan@swsc.com.cn
	杨薇	销售岗	15652285702	15652285702	yangwei@swsc.com.cn
	姚航	销售岗	15652026677	15652026677	yhang@swsc.com.cn
	张鑫	销售岗	15981953220	15981953220	zhxin@swsc.com.cn
王一菲	销售岗	18040060359	18040060359	wyf@swsc.com.cn	

	王宇飞	销售岗	18500981866	18500981866	wangyuf@swsc.com
	马冰竹	销售岗	13126590325	13126590325	mbz@swsc.com.cn
	龚之涵	销售岗	15808001926	15808001926	gongzh@swsc.com.cn
	杨举	销售岗	13668255142	13668255142	yangju@swsc.com.cn
广深	唐茜露	销售岗	18680348593	18680348593	txl@swsc.com.cn
	文柳茜	销售岗	13750028702	13750028702	wlq@swsc.com.cn
	林哲睿	销售岗	15602268757	15602268757	lzl@swsc.com.cn
