

AI 赋能资产配置（十七）

AI 盯盘：“9·24”行情案例

◆ 策略研究 · 策略解读

证券分析师：	王开	021-60933132	wangkai8@guosen.com.cn	执证编码：S0980521030001
证券分析师：	陈凯畅	021-60375429	chenkaichang@guosen.com.cn	执证编码：S0980523090002

事项：

金融市场中的短期内快速上涨行情往往因情绪驱动而非基本面改善，容易导致阶段性追高。传统技术分析（如 KDJ、RSI、MACD、均线体系、成交量、换手率及估值水平等单一指标）虽能提供部分洞察，但其信号纷杂、滞后性强且受主观经验影响较大，难以有效预警此类脉冲式行情的风险。

为系统性地解决这一问题，本研究旨在构建一个多维度、量化、由人工智能驱动的综合研判框架。研究首先从趋势、动量、资金流向、估值四个核心维度出发，构建了十二个关键原始指标，形成一个全面刻画市场状态的多因子体系，并初步判断和市场趋势的关系，确保模型的可行性以及避免机器学习的“黑箱”现象。随后，本研究创新性地引入以 HistGradientBoosting 为代表的机器学习算法作为“AI 辅助”核心引擎，在两大关键环节实现优化：

在因子处理层面：利用 AI 方法自动化完成因子的去极值、标准化、分位化与特征重要性筛选，显著提升因子数据的稳定性和有效性，克服了传统人工选择因子的主观性与过拟合缺陷。在模型建模层面：利用 HistGradientBoosting 模型的非线性拟合能力，深度捕捉多个技术因子与估值因子之间复杂的交互效应与协同关系，从而生成更为精准、稳健的复合择时信号。

本研究的核心价值在于解答两个关键问题：第一，在 A 股市场的特定环境下，市场因子对于规避短期顶部风险具有显著的预测能力？第二，相较于传统的线性模型或经验判断，AI 辅助的融合框架能否在策略胜率、夏普比率与最大回撤控制上表现出更优的性能，尤其在于类似“9·24”行情的临界点识别上。

我们预期，该 AI 增强型的多因子择时策略不仅能为投资者提供一个客观、系统的 A 股风险预警工具，减少盲目追高带来的损失，同时也为人工智能技术在中国金融市场量化实践中的应用潜力与可解释性提供了有力的实证参考。

报告首先通过趋势、动量、资金流向和估值四个维度的图表验证了单点指标在市场波动中能够捕捉趋势拐点与风险信号，为量化建模提供了实证基础。在此基础上，研究团队将 12 个核心指标及其衍生因子系统化处理，统一标准化和分位化，并通过 IC 检验筛选出具备解释力的因子。随后引入直方图梯度提升（HistGradientBoosting）模型，利用其非线性拟合与交互捕捉能力生成复合择时信号。回测结果显示，策略年化收益率达 36.41%，夏普比率 2.30，最大回撤-19.51%，显著优于万得全 A 指数，尤其在“9·24”行情等高波动场景中展现出“追涨不追高、下跌能抗跌”的特征。由此证明，AI 辅助的多因子量化框架不仅能将单一指标的局部关联转化为系统化策略，还能有效提升风险控制与收益稳定性。

观点：

◆ 从“9·24”说起，四维度打造市场放大器

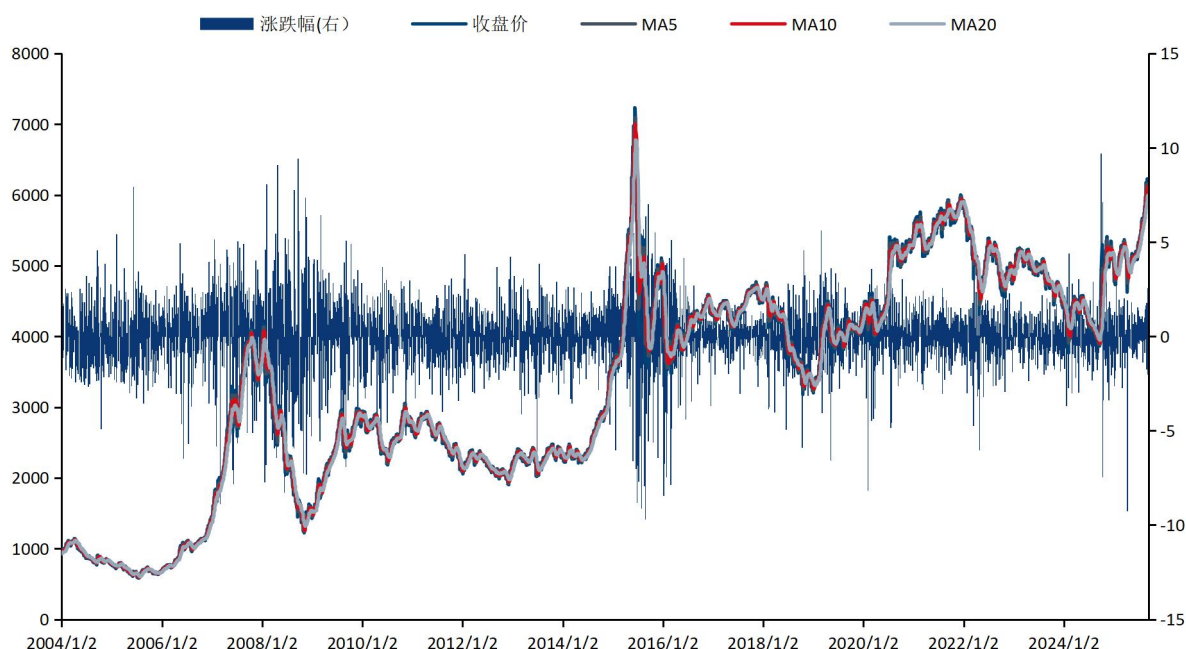
回顾 2024 年 9 月下旬，A 股市场曾经历了一波短暂而迅猛的“9·24”快速上涨行情。9 月 24 日万得全 A 指数收盘于 4079 点，随后在短短四个交易日内连续拉升，于 9 月 30 日收至 4941 点，累计涨幅超过 21%。10 月 8 日更是一度冲高至 5294 点。然而行情并未延续，10 月 9 日起市场快速回落，至 10 月 11 日已跌回 4742 点附近，短期内形成“急涨—急跌”的典型脉冲式行情。这一波动使大量追高资金被套在阶段性高点，

验证了市场在缺乏基本面支撑时情绪驱动上涨的脆弱性与高风险性。这一案例清晰揭示出：仅凭市场情绪推动的短期脉冲上涨，往往缺乏坚实的基本面支撑，其脆弱性极易对趋势交易者造成显著伤害。也因此引发了一个值得深度探讨的核心问题：类似“9·24”行情这种突如其来的短期急涨，市场中哪些指标可以提前反应出来，以及在右侧高频监测风险呢？

趋势角度：趋势启动，趋势延续和震荡回调信号

从数据上看，9月24日万得全A收盘价为4079点，当时MA5>MA10>MA20已经开始显现，均线形成了初步的“多头排列”。到9月30日收盘价升至4941点，MA5、MA10、MA20完全打开，典型的趋势启动信号。在“9·24”行情冲高至10月8日后，市场很快进入震荡回调阶段。具体来看，10月10日开始，MA5已经逐步贴近收盘价，显示短期上行动能减弱；至10月15日至17日，均线关系出现了MA10>MA5>MA20的结构，反映出短期均线被中期均线压制，行情由单边上行转向震荡整理。与此同时，布林带宽度在快速扩张后开始收敛，进一步印证了趋势持续性减弱。整体上，这一均线“反压”与波动收缩的组合，构成了较为清晰的回调预警信号，提示追高资金的风险正在加速累积。

图1：收盘价与均线趋势下的市场波动特征



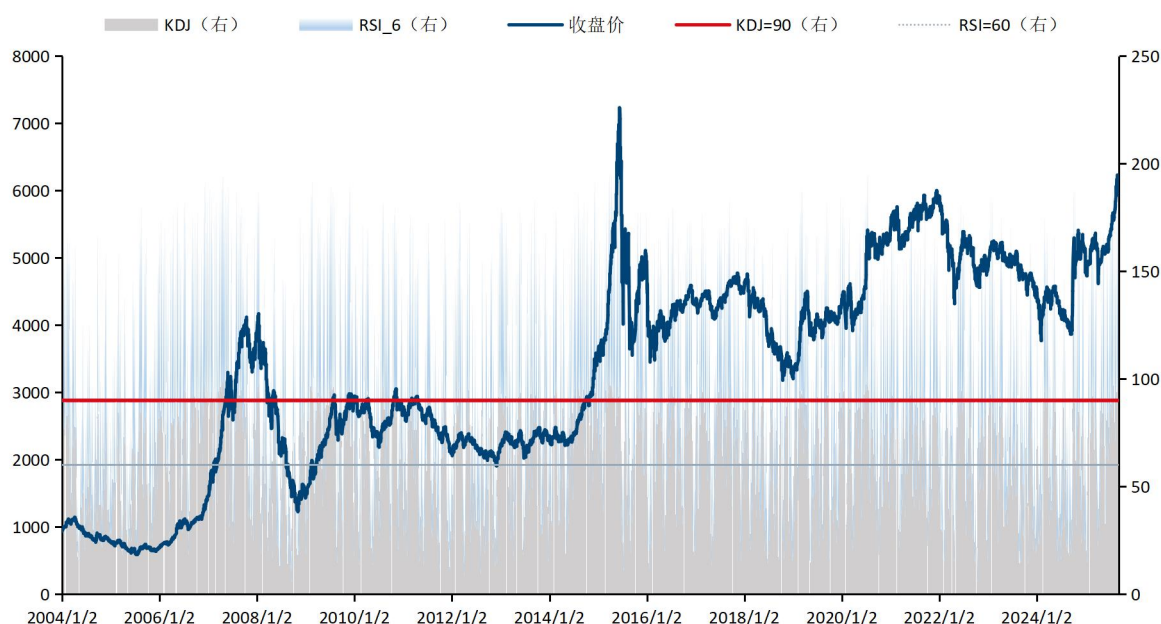
资料来源：万得，国信证券经济研究所整理

动量角度：上涨波段的动量极值，回调启动的动量转折以及波段循环的临界点

在上涨波段中，KDJ和RSI指标同步进入极高区间，往往预示着动量已经进入加速阶段。以“9·24”行情为例，指数从4079点拉升至4941点期间，KDJ升至接近90，RSI_6更是逼近97，显示市场处于典型的超买状态。这种动量极值往往出现在波段尾声，是短期行情见顶的重要信号。随后在行情回调启动阶段，KDJ与RSI快速从高位跌落，动量指标同步转弱。10月9日至11日，随着指数回落至4742点，KDJ从89急降至66，再到56，RSI也从98跌至58，短短数日内跌破关键阈值60。这一动量转折明确释放了行情结束的预警。

从更长的波动循环来看，KDJ在90附近往往对应波段高点，而在20-30附近对应底部；RSI则以60一线为分界，上破60通常意味着新一轮上涨启动，下破60则提示上涨结束。因此，KDJ的极值与RSI的分界线共同构成了波段循环的临界点，有助于投资者在短期急涨急跌行情中把握节奏。

图2: KDJ 与 RSI 阈值验证短期行情见顶信号



资料来源：万得，国信证券经济研究所整理

资金流向角度：成交额与价格的联动和背离

在“9·24”行情中，成交额的起伏与指数的上涨幅度高度相关。9月24日万得全A成交额仅为9744亿，但随着行情启动，9月27日成交额迅速放大至14556亿，9月30日更是飙升至26115亿，与指数从4079点上涨至4941点高度同步。随后10月8日冲顶至5295点，成交额达到34835亿的阶段高点。数据表明，成交额的相对起伏越大，价格的攀升幅度往往也越大。

量能启动先于价格突破：值得注意的是，成交额的爆发往往领先于价格的快速拉升。例如在9月25-27日期间，成交额连续放大至1.16万亿~1.45万亿，而指数则在9月27日之后才加速突破至4500点上方。这说明，量能的提前启动是行情加速的重要前置信号，投资者若能及时捕捉这一变化，往往可以更早发现趋势性机会。

在10月9日之后，指数虽仍在4900-5000点一带震荡，但成交额已明显回落：10月9日成交额从34835亿降至29666亿，10月11日进一步降至15871亿。这种量能萎缩与价格高位震荡形成背离，预示行情的动能减弱，随后指数便快速下探至4742点，验证了资金流出对行情见顶的警示作用。值得一提的是，2025/4/1到5/1出现了一个比较明显的价量背离，这表明成交额先行萎缩，价格仍在惯性冲高，它提示行情缺乏新的资金支撑，涨势难以长久，后期价格果然回调。到了5/6以后，成交额突然放大，才真正配合价格突破。所以4月的背离实际上是一个“警示阶段”，而5月初的补量才是真正确认行情能走远的信号。

图3: 收盘价与成交额的联动与背离关系



资料来源: 万得, 国信证券经济研究所整理

估值角度: 估值抬升预告风险积累, 高分位与技术超买共振增加风险

从 2024/9/24 的 PE(TTM)15.63、5 年分位仅 3.6%, 到 2024/9/30 时 PE 已经攀升至 18.45, 分位数飙升至 65%。进一步在 2024/10/8, PE 达到 19.55、分位 74%, 正好对应行情高点 (5295 点)。这表明, 当估值分位进入 >70% 区间时, 市场已从低估快速转向高估, 风险溢价明显下降。进入年底和次年初, 2024/12/5 和 2025/1/14 的分位数分别为 0.701、0.715, 都对应价格阶段性高点, 并伴随 KDJ 和 RSI 的短期见顶。进一步到 2025/3/20 附近, 分位数再次升至 0.71, 此时市场冲高乏力, 很快进入 4 月的调整期。

整体来看, PE 分位数高于 70% 本身意味着市场估值偏向历史高位, 而当这一信号叠加动量指标的超买 (如 RSI 超过 70、KDJ 在 90 以上), 往往成为短期行情的风险提示。在多个阶段的对比中, 这一组合信号都有效预示了行情的见顶和回调压力, 具备较强的解释力和前瞻性。

图4: 收盘价与 PE 五年分位数的同步性与风险提示



资料来源: 万得, 国信证券经济研究所整理

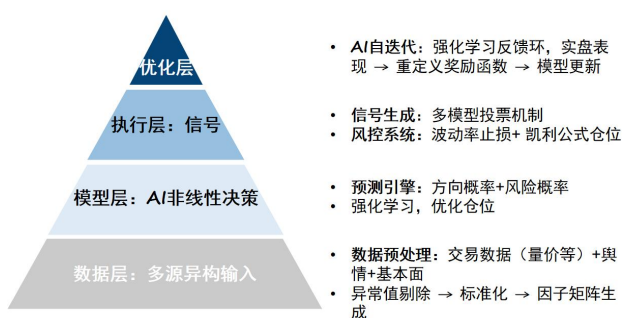
◆ AI 驱动的量价策略流程

AI 驱动的量价策略“全自动引擎”由四个核心层次构成。数据层汇集多源异构输入，包括交易数据、估值指标及基本面因子，并通过异常值剔除与标准化处理生成因子矩阵；模型层则基于 AI 非线性决策方法，结合方向概率预测与风险框架，利用机器学习和优化算法形成可执行的仓位方案；执行层通过多模型投票机制生成信号，并与波动率止损、杠杆与仓位控制等风控系统相结合，实现策略落地；在此之上，优化层通过强化学习反馈环与实盘表现评估，不断对奖励函数与模型进行迭代更新。整体流程形成了从数据到信号、再到执行与优化的闭环，实现了量价择时的自动化与自我进化。

特征工程是 AI 量价策略的关键前置环节。首先多源数据出发，经过清洗环节对缺失值进行填补，并通过异常值处理与标准化手段确保数据的一致性与可比性；随后，将数据按时间顺序切分为训练集、验证集和测试集。接着，模型根据预处理后的数据生成各类特征，包括技术指标、量价特征以及风险因子等，并整合为特征矩阵。在此基础上，还需经过防泄漏验证，确保特征构造不引入未来信息或数据穿越。最终得到的特征集能够为后续建模与策略优化提供坚实的数据基础。

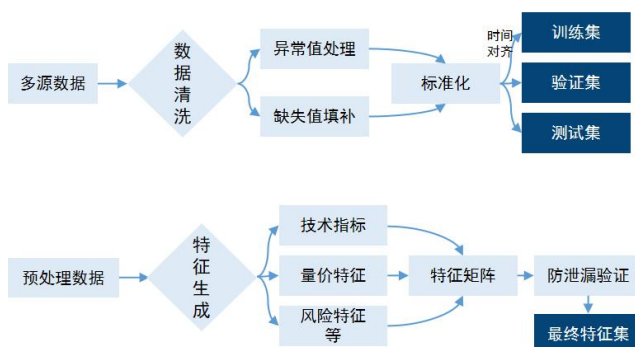
整体来看，AI 驱动的全自动引擎为量价择时提供了完整的框架，但要真正发挥其效能，仍需从底层指标体系出发，确保输入因子具备系统性和解释力。这也正是我们从单点指标迈向量化框架、构建原始与衍生指标体系的出发点。

图5: AI 驱动的量价策略“全自动引擎”



资料来源：万得，国信证券经济研究所整理

图6: 特征工程示意图



资料来源：万得，国信证券经济研究所整理

◆ 从单点指标到量化框架：原始/衍生指标体系构建

虽然趋势、动量、资金和估值四个角度的指标都能反映市场运行的某些特征，但单独依赖这些指标去判断市场方向，往往存在滞后性和片面性，难以真正实现风险规避。例如，均线交叉可能在行情启动后才确认，动量指标容易在短期波动中频繁发出“假信号”，资金流向和估值分位虽然能揭示结构性风险，但也无法精准捕捉转折点。因此，仅凭这些单点指标与市场趋势的经验性关联，还不足以形成系统化的判断框架。

基于这一局限，我们需要进一步迈向量化策略的搭建。第一步，就是在公开数据库中从四个维度精选出一组代表性指标，既涵盖趋势类（MACD、MA5/MA10/MA20）、动量类（KDJ、RSI_6）、资金与市场活跃度类（成交额、换手率、振幅、涨跌幅），也包括估值类（PETTM 及其分位数），再辅以收盘价作为基准。这样，我们就形成了 12 个核心指标池，为后续的因子处理、有效性检验以及策略构建奠定基础。这些因子具有较强的市场解释力，并在量化研究与实务交易中得到广泛应用，同时为了让指标具有更好的解释力度，我们把原始指标转换为一些衍生指标：

表1: 12 个核心指标池

指标名称	含义	计算方式/说明
MACD	指数平滑异同移动平均线，用于刻画价格趋势及其强弱	由长期与短期指数移动平均线（EMA）的差值构成，并结合信号线判断趋势
KDJ	随机指标，反映价格短期超买超卖情况	基于最高价、最低价与收盘价计算随机值 K、D、J

RSI_6	相对强弱指数（6 日），衡量价格动量	$RSI=100-100/(1+RS)$ ，其中 RS 为 6 日平均上涨幅度与下跌幅度之比
收盘价	每日市场交易收盘价，作为价格基准	直接取交易日收盘价
涨跌幅	每日收益率，反映价格相对变化幅度	$(\text{当日收盘价}-\text{前一日收盘价})/\text{前一日收盘价}$
成交额	市场交易活跃度的重要量能指标	当日成交量×当日成交均价
振幅	日内波动区间与价格关系，反映市场活跃度	$(\text{当日最高价}-\text{最低价})/\text{前一日收盘价}$
PE(TTM)	滚动市盈率，衡量估值水平	市值/最近 12 个月归母净利润
MA5/MA10/MA20	不同周期的移动平均线，反映趋势	$MA_n=\text{收盘价的 } n \text{ 日算术平均}$
换手率	市场资金流动与交易活跃程度的核心指标	当日成交量/流通股数

资料来源：万得，国信证券经济研究所整理

指数平滑异同移动平均线（MACD）是趋势分析中最常用的指标之一。它通过短期与长期指数移动平均线的差值来判断趋势的方向与力度，配合信号线的交叉，可识别趋势的启动与反转。相比原始 MACD 值，加速度和偏离度能够提供更多关于趋势变化速率与异常状态的信息，有助于捕捉临近的趋势转折点。因此，我们构建了：

表2: MACD 衍生指标

衍生因子名称	含义/计算方式	作用与逻辑
MACD_acc	MACD 的一阶差分，即当日 MACD 值-前一日 MACD 值	用于识别趋势加速或减缓，捕捉趋势变化速率
MACD_anom	MACD 与其历史均值的偏离程度	衡量趋势的异常性或过热状态，提示潜在反转或修正

资料来源：万得，国信证券经济研究所整理

KDJ：随机指标，通过价格在一定区间内的相对位置来衡量超买超卖状态。数值偏高通常表示短期内可能存在超买风险，偏低则提示超卖机会，适合捕捉短期反转。

RSI_6：相对强弱指数的短周期版本（6 日），主要反映价格的动量变化。其计算基于上涨幅度与下跌幅度的平均值之比，数值越高代表上涨动能越强。因此我们进行了分位化处理得到 RSI_percentile63：即 RSI 在过去 63 日内的分位水平。通过滚动分位数替代绝对阈值，增强了 RSI 在不同波动阶段的适应性，更合理地判断市场动量状态。

收盘价：每日交易结束时的价格，是市场最重要的基准价格。收盘价不仅直接用于收益率计算，也为均线、突破类指标提供计算基础。收盘价本身只提供点位信息，而通过均线、突破和斜率等二次加工，可以更全面地刻画趋势的方向性与强弱：

表3: 收盘价衍生指标

衍生因子名称	含义/计算方式	逻辑
Break20	判断收盘价是否突破过去 20 日高点（收盘价>20 日最高价）	识别趋势启动信号
Slope10	对最近 10 日收盘价进行线性回归，取回归斜率	衡量短期趋势方向与强弱

资料来源：万得，国信证券经济研究所整理

涨跌幅：反映资产相对价格变化幅度的核心指标，通常以百分比形式表示。它是构建收益率序列和波动率的基础，也是金融建模的关键输入。

成交额：代表市场在某一交易日的资金参与程度。成交额的放大通常意味着资金入场活跃，反之则可能预示市场观望情绪，是市场活跃度和资金流动的直接体现，但原始值容易受市场规模变化影响。因此对其进行标准化和边际变化处理，成交额均值反映常态水平，边际变化与爆发性信号则能捕捉资金流入的关键节点，往往是趋势启动的确认信号：

表4: 成交额衍生指标

衍生因子名称	含义/计算方式	作用与逻辑
MA10_VAL/MA20_VAL	成交额的 10 日、20 日滚动均值	作为市场交易活动的基准水平, 衡量常态成交水平
$\Delta \text{Value10}/\Delta \text{Value20}$	(当前成交额-MA10/20_VAL)/MA10/20_VAL	刻画资金相对常态的边际放量或缩量, 反映资金流动变化
Burst10/Burst20	当前成交额/MA10/20_VAL	识别成交额的爆发性增长, 捕捉资金突然放大信号

资料来源: 万得, 国信证券经济研究所整理

振幅: 衡量日内价格波动区间相对前一日收盘价的大小。振幅较高时, 往往意味着市场不确定性增强。为了消除不同标的和不同时期振幅的不可比性, 对其进行了标准化。

PE(TTM): 滚动市盈率 (过去 12 个月 EPS 为基准), 用于估值分析。高估值区间可能意味着透支未来预期, 低估值区间则可能具备吸引力。为了避免绝对数值不可比, 对 PE 进行了历史分位处理, 得到 PE_pct_5y, 即 PE 在过去 5 年中的百分位水平。

MA5、MA10、MA20: 不同时间尺度的移动平均线。短期均线 (MA5) 反映市场即时趋势, 中期均线 (MA10/MA20) 则用于识别趋势的持续性与拐点。

表5: 移动均线衍生指标

衍生因子名称	含义/计算方式	逻辑
Price_MA5_gap	(收盘价-MA5)/MA5	判断短期超买/超卖
Price_MA10_gap	(收盘价-MA10)/MA10	判断中期超买/超卖
Price_MA20_gap	(收盘价-MA20)/MA20	判断中期均值回归偏离程度
MA_channel_w	(MA20-MA5)/MA20	刻画均线通道宽度, 反映价格波动区间大小
sign_5_10	MA5 与 MA10 交叉信号 (金叉=1, 死叉=-1)	常用趋势反转指标
交叉年龄	最近一次 MA5 与 MA10 交叉以来的交易日数	衡量趋势持续时间

资料来源: 万得, 国信证券经济研究所整理

换手率: 衡量市场流动性与交易活跃度的重要指标。换手率高说明股票在市场中的流动性强, 投资者参与度高。我们同样进行了标准化处理, 得到 Turnover_z20, 即换手率的 20 日 Z-score。

布林带: 布林带是一种常见的波动率通道指标, 包含价格位置和波动宽度两类信息。我们试图构建布林带相对位置和波动状态的衍生指标来识别趋势突破, 和反映市场的收敛-扩张循环。

表6: 布林带衍生指标

衍生因子名称	含义/计算方式	作用与逻辑
BB_percentB	(收盘价-下轨)/(上轨-下轨)	衡量收盘价在布林带中的相对位置, 识别价格高低位
BB_bandwidth	(上轨-下轨)/MA20	表示布林带宽度, 刻画波动水平大小
BB_bw_pct	布林带宽度在历史序列中的分位水平	比较当前波动与历史波动状态, 识别极端收缩或扩张
BB_squeeze_flag	当带宽低于历史低分位 (如 20%分位) 时取 1, 否则为 0	捕捉布林带收缩信号, 往往意味着行情即将选择方向
BB_expand_start	当布林带宽度由低位向上突破时取 1	表征布林带扩张起点, 代表行情波动显著放大阶段

资料来源: 万得, 国信证券经济研究所整理

◆ 因子处理: 面向实战的标准化与分位变换

在完成原始/衍生指标体系构建后, 下一步的关键环节是因子化处理。具体流程是: 首先, 将选取的指标序列转换为统一方向的因子; 其次, 对因子进行去极值、标准化或分位化, 消除不同口径和量纲间的差异; 最后, 将处理后的因子作为输入, 进入因子解释力度检验和机器学习建模。在因子处理环节, 本研究主要采用两类方法: Z-score 标准化和 GaussianRank (高斯分位)。

Z-score 标准化: 通过将因子值减去均值并除以标准差, 将数据转化为均值 0、标准差 1 的分布。其含义是衡量单个数据点相对于总体均值的偏离程度 (以标准差为单位)。正值表示高于均值, 负值表示低于均值, 绝对值越大偏离越远。

GaussianRank 变换：也称为分位数归一化到正态分布。其原理是先对因子值进行排序，以其在样本中的百分位替代原始值，然后将这些百分位按照逆正态累积分布函数映射为对应的高斯分布值。简单来说，就是将因子值按从小到大 rank，赋予均匀分布的分位值，再映射到标准正态分布上（均值 0、标准差 1）。

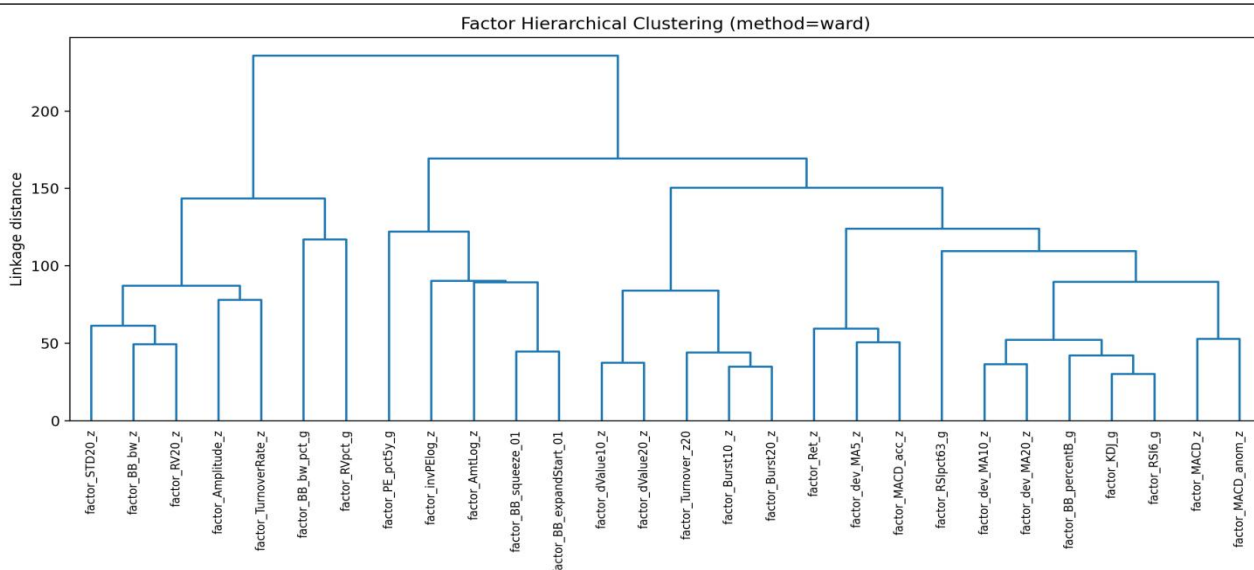
图7：因子处理方法与标准化结果示例

Z-score	Gaussian Rank
factor_KDJ_g	factor_MACD_z
factor_RSI6_g	factor_dev_MA5_z / 10_z / 20_z
factor_RSIpct63_g	factor_Ret_z
factor_RVpct_g	factor_STD20_z
factor_PE_pct5y_g 等	factor_Amplitude_z
	factor_BB_bw_z
	factor_RV20_z
	factor_AmtLog_z、
	factor_TurnoverRate_z等

资料来源：万得，国信证券经济研究所绘制

通过因子处理，原始指标不再是杂乱无章、口径各异的数列，而是被转化为具备可比性、可聚合性和可回测性的交易因子。一方面，标准化和秩变换有效消除了不同量纲之间的差异，使成交额、估值分位与趋势偏离等数据能够在统一尺度下比较；另一方面，GaussianRank 等方法增强了因子的稳健性，削弱极端值的干扰，避免“虚假超买/超卖”信号带来的误导。更重要的是，因子处理为策略研究带来了可复现性与可落地性，使得这些因子能够直接用于回测、组合构建和机器学习建模，从而实现对市场状态的系统化评估，并为投资者提供更科学的择时与风险管理依据。

图8：因子层次聚类结果（Ward 方法）



资料来源：万得，国信证券经济研究所绘制

◆ AI 辅助模型选择

在完成因子体系的标准化与构建之后，下一步的核心是如何将这些因子有效组合并转化为可操作的交易信号。传统方法通常依赖线性模型或经验加权，但在面对数量庞大、分布复杂且相互关联的因子时，容易出现信息损失或过度简化，难以捕捉非线性关系和交互效应。为此，本研究引入人工智能辅助模型，以机器学习算法作为因子整合与择时信号生成的引擎。其优势在于：一方面能够自动识别高维因子间的非线性与非平稳关系，另一方面通过训练与验证提升信号的稳健性和泛化能力，从而在多因子体系的基础上实现更高效的择时策略构建。

我们直接通过 IC 值检验与有效性筛选来判断并保留有解释力的因子。这一方式不仅保证了模型的透明度与可解释性，也避免了“黑箱化”处理，使得研究结果更易于理解与应用。

表7: 常见机器学习模型及其参数与应用场景

名字	介绍	主要参数	使用场景
线性回归(LinearRegression)	最基础的回归模型，假设输入与输出呈线性关系	正则化系数(L1/L2)、学习率	因子数量：少量 (<50)，需低相关；处理方式：标准化、去共线；金融应用：收益率预测、宏观因子回归
逻辑回归(LogisticRegression)	分类模型，输出类别概率，本质是广义线性模型	正则化系数、迭代次数	因子数量：几十个以内；处理方式：标准化、分箱 (WOE 转换)；金融应用：信用评分、二分类买卖点
决策树(DecisionTree)	基于特征划分的树状模型，结果直观可解释	最大深度、最小样本数、分裂标准	因子数量：10-100；处理方式：无需标准化；金融应用：贷款审批、投资者风险分层
随机森林(RandomForest)	多棵树投票/平均，降低过拟合	树的数量、最大特征数、树深度	因子数量：几十到几百；处理方式：无需标准化，可自动处理缺失；金融应用：因子选股、组合风险预测
梯度提升树 (GBDT/XGBoost/LightGBM/CatBoost)	迭代训练弱分类器，精度高	学习率、树深度、迭代次数、正则化参数	因子数量：上百个，支持稠密/稀疏特征；处理方式：缺失值自动处理，分桶稳定；金融应用：择时模型、成交率预测
直方图梯度提升 (HistGradientBoosting,HGB)	GBDT 的高效实现，直方图近似加速，适合中等规模因子和中大样本量	学习率（步长大小）、max_bins（直方图分桶数）、树深度（模型复杂度）、最小样本数	因子数量：20-100 个最合适（例如 28 个择时因子）；处理方式：自动分桶，减少噪声；无需标准化；金融应用：量化择时策略（因子合成信号→预测未来涨跌）、大规模表格数据因子建模、智能投顾推荐
神经网络 (NeuralNetworks/DeepLearning)	多层非线性结构，适应复杂数据	网络层数、神经元数、激活函数、学习率	因子数量：几百-几千；处理方式：标准化，必要时 PCA/Dropout 降维；金融应用：高频交易信号、波动率建模
隐马尔可夫模型(HMM)	基于隐状态的序列建模	状态数、转移/发射概率矩阵	因子数量：少量时序因子（收益率、波动率）；处理方式：序列平稳化；金融应用：牛熊市切换、市场状态建模

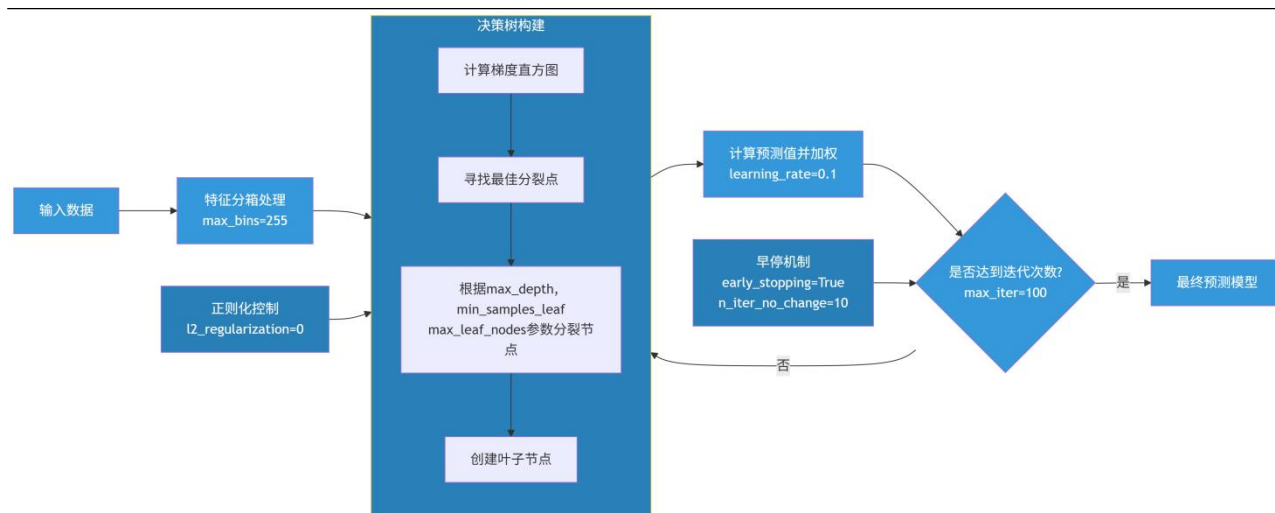
资料来源：万得，国信证券经济研究所整理

经过因子筛选和标准化处理，我们需要一个模型把这些因子转换为具体的交易信号。在常见的机器学习模型中，我们最终选择了 HistGradientBoosting（直方图梯度提升）。它本质上是改进过的树模型，既能处理非线性关系，又在速度和效果之间取得了很好的平衡，非常适合量化策略。

在参数设定上，我们保持了一个比较实用的调参逻辑。先用较大的学习率（0.1）配合不同的树深度（max_depth）测试，找到既能捕捉规律又不会过拟合的复杂度。然后再在合适的深度下，通过降低学习率、增加迭代次数（max_iter）来让模型更稳健。训练过程中，数据按照 7:3 的比例切分为训练集和验证集，我们重点观察验证集的表现，比如夏普比率、未来收益的相关性（IC）和回撤情况。如果深度过大虽然收益提升，但回撤明显放大，就会优先选择更稳健的配置。

模型训练完成后，会输出一个“看涨概率”分数。我们将其直接转化为交易信号：如果未来 5 日收益率上涨的概率超过 60%，就发出买入信号，否则保持观望或卖出。这种做法既避免了过度复杂的建模，又能保证策略在实际运行中的可解释性和可操作性。

图9：直方图梯度提升（HGB）模型的训练流程



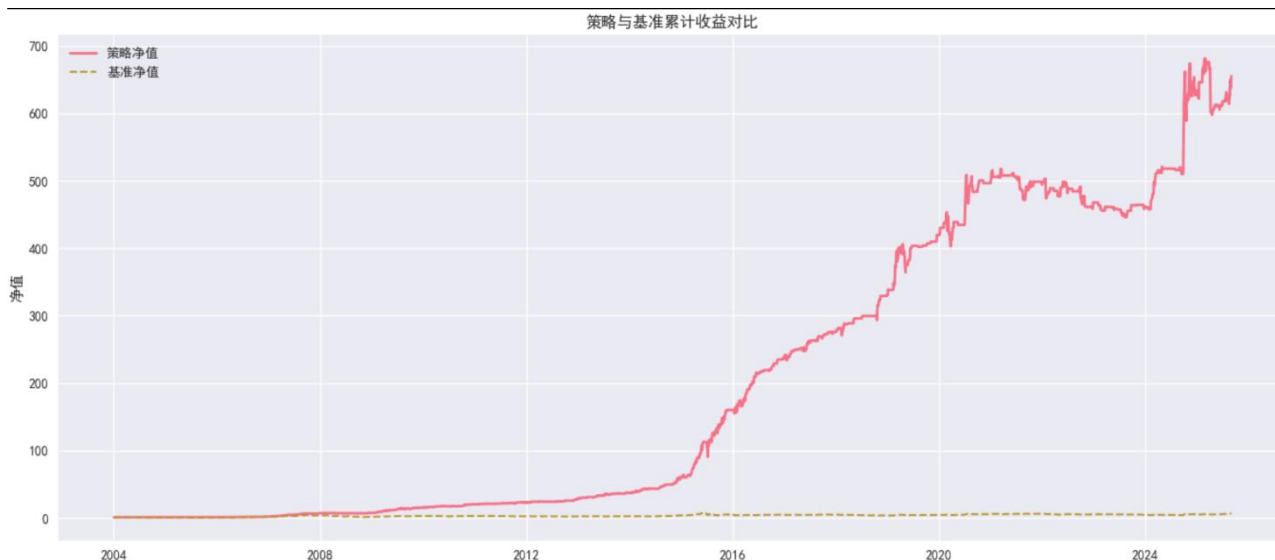
资料来源：万得，国信证券经济研究所绘制

◆ 量化择时策略效果分析

模型训练完成后，我们将其应用于历史数据进行回测，评估策略的实际表现。整体来看，策略的累计净值曲线显著优于万得全 A 指数，长期呈现稳健上行特征，并且在大多数市场下跌阶段表现出较强的抗跌性。基准万得全 A 指数在样本期内波动剧烈，经历多次大幅调整，而策略曲线在这些阶段往往能够有效控制回撤，显示出良好的风险管理能力。整体回测结果表明，AI 盯盘的策略模拟盘年化收益率约 36.41%，年化波动率维持在 15.81% 左右，夏普比率达到 2.30，明显高于市场基准。

从阶段表现看，策略在多头牛市中均能紧跟上涨趋势并获得超额收益，而在市场大跌时通过及时减仓或发出卖出信号有效回避了深度下行。例如，在 2024 年 9 月 24 日（“9·24”行情）前后，万得全 A 自 3916 点快速拉升至 4941 点，短短一周涨幅超过 25%，期间成交量快速放大，市场情绪极度亢奋。然而，策略并未在初始阶段盲目加仓，而是等到信号强烈确认后逐步跟进，并在 9 月 30 日左右净值迅速突破 57 万（相比行情启动前的 47 万大幅增长）。当指数在 10 月 9 日至 10 月 17 日快速回调约 12% 时，策略净值仅小幅回落，基本守住了大部分涨幅。这说明模型在捕捉快速上涨行情的同时，也能在高波动期减少损失，具备良好的“追涨不追高、下跌能抗跌”的特征。

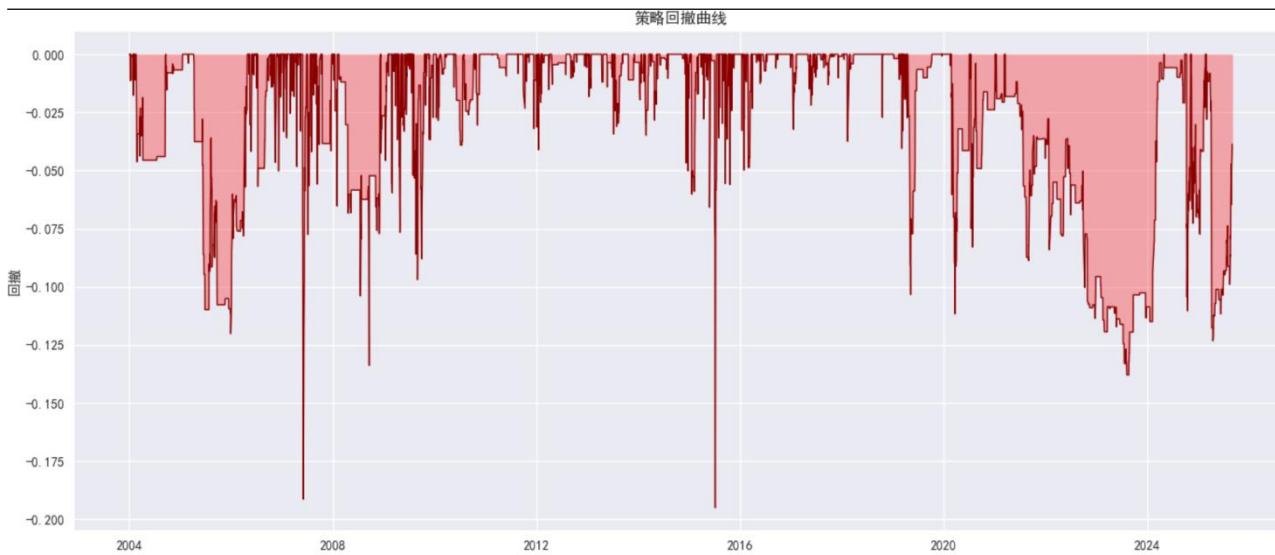
图 10: AI 择时策略回测表现与基准比较



资料来源：万得，国信证券经济研究所整理

在评估策略表现时，最大回撤是衡量风险控制能力的重要指标。从回测结果来看，策略整体回撤水平较为可控，最大回撤维持在-19.51%，显著低于万得全 A 同期的深度回撤幅度。首先，在 2004 年初的市场波动阶段，策略虽然出现了几次小幅回撤（如 1 月 9 日的-1.13%、1 月 30 日的-1.77%），但很快就恢复至新高，显示出良好的修复能力。其次，在 2024 年 9 月 24 日（“9·24”行情）期间，指数短期急涨急跌，策略的表现尤为突出。行情启动前策略保持观望，避免了无效波动；行情爆发后逐步跟进，并在 10 月 9 日至 10 月 17 日指数快速下跌约 12% 时，策略的回撤控制在-11% 左右，远小于指数跌幅，表明模型能在极端波动中保护已有收益。总体而言，策略的回撤管理较好、修复速度较快、极端行情下抗跌性强。

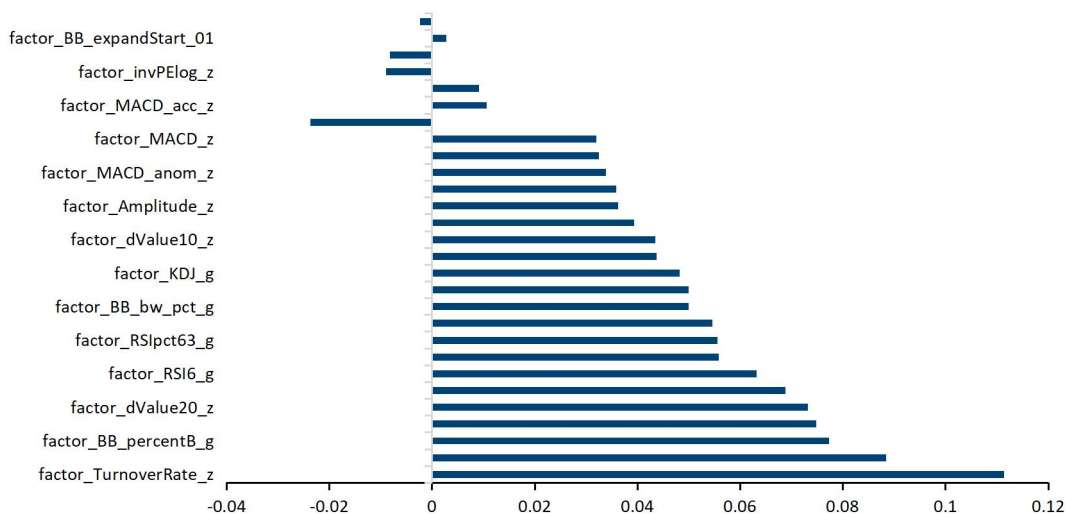
图 11: AI 盯盘策略的回撤曲线示例



资料来源：万得，国信证券经济研究所整理

在构建量化策略时，因子的解释力度和稳定性直接决定了模型预测的有效性。为了验证前期筛选的指标能否真正捕捉市场特征，本研究对所有因子进行了 IC（InformationCoefficient）与 t 值检验，并设定 IC 大于 0.03 且 t 值显著作为有效因子标准。回测结果表明，在全部 28 个因子中，有 21 个因子通过了显著性检验，说明这些因子在长期区间内对未来收益具备稳定的解释力。

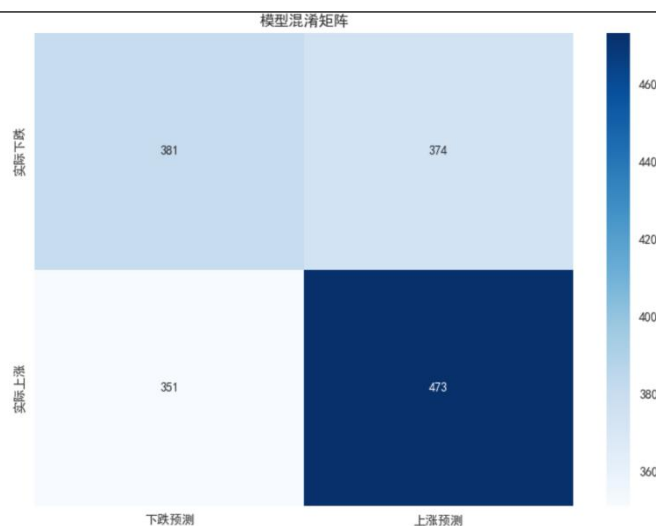
图 12: 通过 IC 值判定有效因子



资料来源：万得，国信证券经济研究所整理

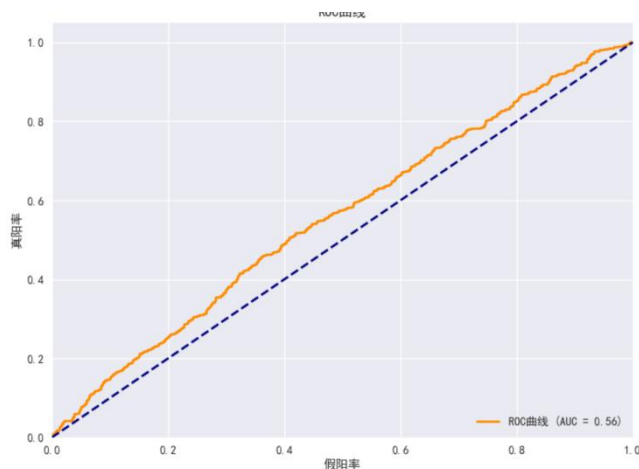
在训练过程中，我们同时计算了混淆矩阵和分类报告来评估模型的预测能力。结果显示：模型在训练集上的准确率约为 74.78%，而在测试集上的准确率为 54.08%。具体来看，在测试集中，模型对上涨情形的识别率（recall）为 57%，对下跌情形的识别率为 50%，总体精确率（precision）维持在 52%–56%之间。这一结果表明，虽然模型在测试集上的表现仅略高于随机猜测（50%），但已经展现出一定的预测能力，并且在捕捉上涨信号方面略占优势。相比于单纯依赖人工经验或静态指标，模型的确学到了一些能够区分市场方向的有效模式。

图 13: 模型预测结果的混淆矩阵



资料来源：万得，国信证券经济研究所整理

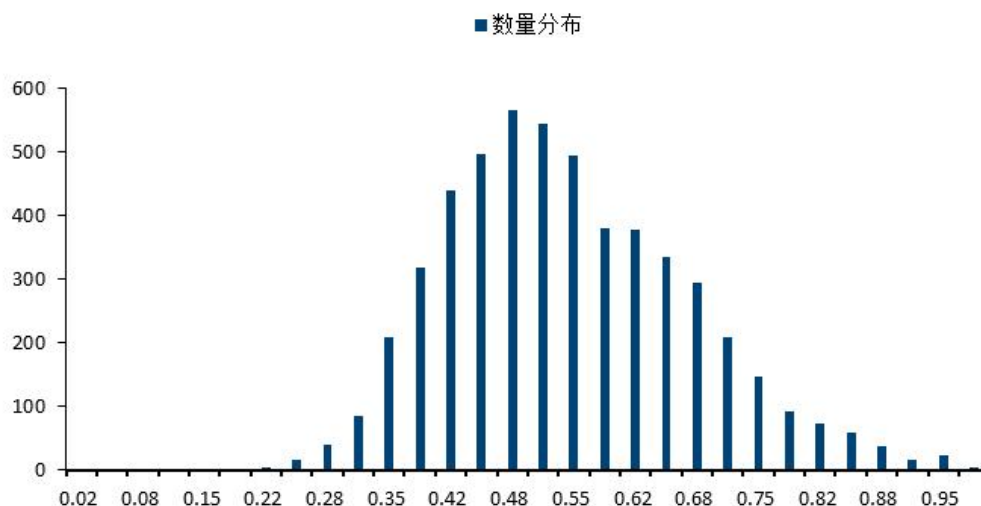
图 14: 模型预测效果的 ROC 曲线与 AUC 值



资料来源：万得，国信证券经济研究所整理

从上涨概率的分布来看，模型输出主要集中在 0.35–0.65 区间，其中 0.45–0.55 这一带占比最高，说明模型在大多数样本上给出的预测概率偏中性，倾向于谨慎判断。这种分布在整体形态上接近正态分布，反映出市场涨跌本身具备较强的对称性和随机性，模型在这一点上并未过度偏向单边行情。一方面，模型的输出具备统计意义上的合理性，没有明显的极端偏差；另一方面，模型对极端行情的捕捉能力仍然不足，可能导致在快速单边行情中反应偏慢。

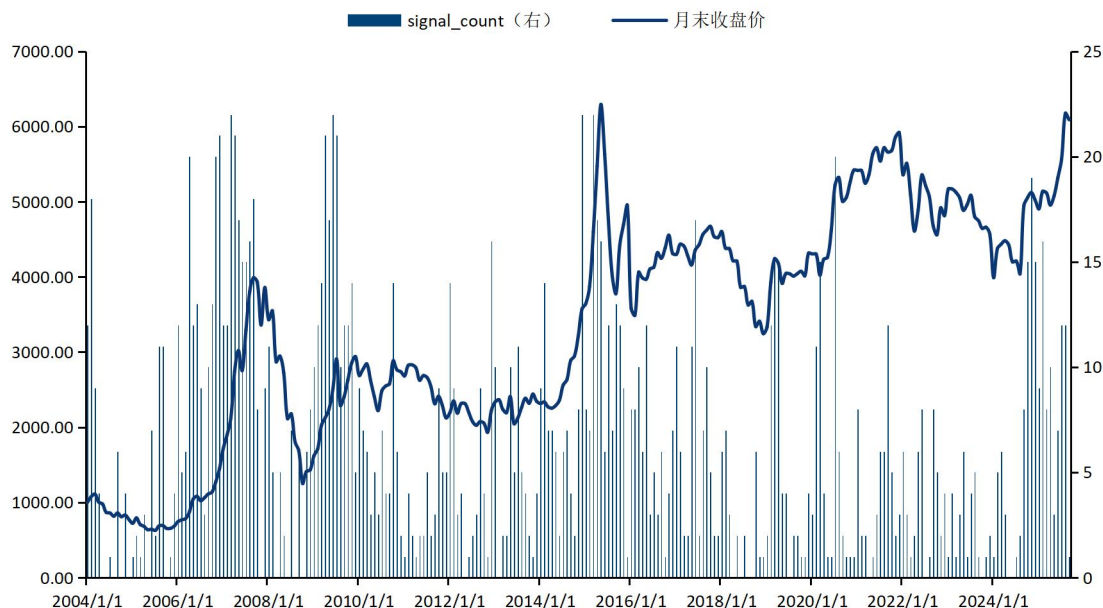
图 15: 模型预测的上涨概率分布直方图



资料来源：万得，国信证券经济研究所整理

从长期统计结果来看，模型在月度层面的信号数量与市场走势之间存在明显的正相关关系：在价格上升阶段，月度信号数量显著增多；而在震荡或下跌阶段，信号数量则相对收敛。这一规律说明模型能够在趋势性行情中捕捉到更多有效的买入或看涨信号，反映了其对价格上升趋势的敏感性与跟随性。尤其是在几轮典型牛市，如 2007 年、2015 年和 2020 年后半段，这意味着模型并非仅仅输出随机噪声，而是在市场上涨周期中确实提供了具有参考价值的择时信号。

图16: 择时信号数量与市场走势的关系



资料来源：万得，国信证券经济研究所整理

综合累计收益曲线和各项指标，我们认为本策略在历史回测中是有效的。它成功利用多因子模型实现了对大盘指数的择时，获取了显著超额收益且控制了下行风险。这证明了从 0 到 1 构建 AI 择时策略的思路是可行的：通过严谨的数据处理、因子筛选、聚类降维，然后选择合适的机器学习模型并精调参数，机器可以从历史中学习市场的某些可预测模式，并转化为实际投资收益。

综上所述，本研究以“9·24”行情为典型案例，系统回顾了 A 股市场中短期情绪驱动型脉冲行情的特征，并提出了一个融合趋势、动量、资金与估值四大维度的多因子量化择时框架。本研究的核心结论在于：AI 辅助的多因子择时框架能够在 A 股市场复杂的波动环境中提供更为系统化、可解释且具有超额收益潜力的风险管理工具。未来，随着特征库的扩展与模型方法的迭代，该框架仍有望进一步提升对极端行情的识别能力，并在实盘应用中不断增强稳健性与适用性。

当前市场确已进入右侧区间，指数位置较去年底显著抬升。以 2025 年 9 月 24 日收盘为例，万得全 A 指数达到 6289.68 点，对应滚动市盈率 22.32 倍，显著高于年初水平，但仍未触及历史极端高位。量能方面，当日成交额保持在约 2.35 万亿元的活跃区间，显示资金流入仍具延续性。换言之，短期虽存在高位震荡的可能，但并未出现全面透支预期的迹象。

中长期维度上，政策环境与流动性条件为市场提供了有力支撑，依然指向乐观行情。国内宏观政策仍强调稳增长与结构转型，资金成本维持低位，上市公司盈利在周期与产业升级中展现韧性；海外方面，美联储已于 9 月开启降息周期，释放出全球流动性转向宽松的积极信号，边际压力趋缓，有助于新兴市场估值修复和资金流入。结合 AI 辅助的量化择时结果，当前市场虽处于右侧运行，但尚未抵达估值与情绪的危险区间，未来仍具备向上空间。整体判断是：短期谨慎对待波动，中长期维持乐观展望，A 股有望在国内政策与海外流动性共振下逐步走出更稳健的上行通道。

◆ **风险提示：**上市企业盈利波动较大；美联储货币政策不确定性；海外地缘政治冲突加剧

相关研究报告：

《策略解读-再议存款搬家》 ——2025-09-06

《策略解读-交易资金眼中的牛市》 ——2025-09-02

《策略解读-时隔半年，再次看多黄金》 ——2025-08-30

《策略快评-2025 年 9 月各行业金股推荐汇总》 ——2025-08-28

《策略解读-科创牛：规律和展望》 ——2025-08-28

免责声明

分析师声明

作者保证报告所采用的数据均来自合规渠道；分析逻辑基于作者的职业理解，通过合理判断并得出结论，力求独立、客观、公正，结论不受任何第三方的授意或影响；作者在过去、现在或未来未就其研究报告所提供的具体建议或所表述的意见直接或间接收取任何报酬，特此声明。

国信证券投资评级

投资评级标准	类别	级别	说明
报告中投资建议所涉及的评级（如有）分为股票评级和行业评级（另有说明的除外）。评级标准为报告发布日后 6 到 12 个月内的相对市场表现，也即报告发布日后的 6 到 12 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。A 股市场以沪深 300 指数（000300.SH）作为基准；新三板市场以三板成指（899001.CSI）为基准；香港市场以恒生指数（HSI.HI）作为基准；美国市场以标普 500 指数（SPX.GI）或纳斯达克指数（IXIC.GI）为基准。	股票投资评级	优于大市	股价表现优于市场代表性指数 10%以上
		中性	股价表现介于市场代表性指数±10%之间
		弱于大市	股价表现弱于市场代表性指数 10%以上
		无评级	股价与市场代表性指数相比无明确观点
	行业投资评级	优于大市	行业指数表现优于市场代表性指数 10%以上
		中性	行业指数表现介于市场代表性指数±10%之间
		弱于大市	行业指数表现弱于市场代表性指数 10%以上

重要声明

本报告由国信证券股份有限公司（已具备中国证监会许可的证券投资咨询业务资格）制作；报告版权归国信证券股份有限公司

关本报告的摘要或节选都不代表本报告正式完整的观点，一切须以我公司向客户发布的本报告完整版本为准。

本报告基于已公开的资料或信息撰写，但我公司不保证该资料及信息的完整性、准确性。本报告所载的信息、资料、建议及推测仅反映我公司于本报告公开发布当日的判断，在不同时期，我公司可能撰写并发布与本报告所载资料、建议及推测不一致的报告。我公司不保证本报告所含信息及资料处于最新状态；我公司可能随时补充、更新和修订有关信息及资料，投资者应当自行关注相关更新和修订内容。我公司或关联机构可能会持有本报告中所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、财务顾问或金融产品等相关服务。本公司的资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告意见或建议不一致的投资决策。

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。任何形式的分享证券投资收益或者分担证券投资损失的书面或口头承诺均为无效。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，我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。

证券投资咨询业务的说明

本公司具备中国证监会核准的证券投资咨询业务资格。证券投资咨询，是指从事证券投资咨询业务的机构及其投资咨询人员以下列形式为证券投资人或者客户提供证券投资分析、预测或者建议等直接或者间接有偿咨询服务的活动：接受投资人或者客户委托，提供证券投资咨询服务；举办有关证券投资咨询的讲座、报告会、分析会等；在报刊上发表证券投资咨询的文章、评论、报告，以及通过电台、电视台等公众传播媒体提供证券投资咨询服务；通过电话、传真、电脑网络等电信设备系统，提供证券投资咨询服务；中国证监会认定的其他形式。

发布证券研究报告是证券投资咨询业务的一种基本形式，指证券公司、证券投资咨询机构对证券及证券相关产品的价值、市场走势或者相关影响因素进行分析，形成证券估值、投资评级等投资分析意见，制作证券研究报告，并向客户发布的行为。

国信证券经济研究所

深圳

深圳市福田区福华一路 125 号国信金融大厦 36 层

邮编：518046 总机：0755-82130833

上海

上海浦东民生路 1199 弄证大五道口广场 1 号楼 12 层

邮编：200135

北京

北京西城区金融大街兴盛街 6 号国信证券 9 层

邮编：100032