基于随机森林（Random Forest）的动态资产配置策略

摘要：

本文提出并比较了两套随机森林驱动的宏观动态资产配置策略：1）仅预测“未来12个月负收益概率P”的单目标随机森林模型策略；2）同时预测“未来12个月深度回撤概率K”与“未来6个月正收益概率R”的双目标随机森林模型策略。研究首先对 2010-2025 年中国月度宏观数据做滚动Z-score标准化，并按“增长-通胀-流动性”进行宏观因子的分类聚合；随后在滚动窗口内采用SMOTE-ENN+随机森林模型训练分类器，并将概率连续映射为股票、国债、现金、大宗商品、黄金五类资产权重（按月调仓、设最小阈值抑制换手）。本文回测结果：单目标策略取年化收益6.82%、波动率11.24%、最大回撤−25.3%，Sharpe为0.64；双目标策略进一步全面优化至 7.05%/8.91%/−14.9%/0.81，两者均显著优于静态60/40基准策略（5.77%/17.2%/−37.3%/0.41）。我们证明了利用宏观信息经机器学习量化并实时调节权重，能够在多轮牛熊切换中获得更优的风险调整后收益，且双目标框架在进攻-防守平衡方面表现最佳。

**关键词**：宏观经济因子；随机森林；动态资产配置

1. 引言
   1. 研究背景与动机

随着全球经济的日益复杂和金融市场波动的加剧，传统资产配置策略（如均值-方差方法）因其依赖于资产历史表现的稳定性假设，难以适应迅速变化的市场环境，常常表现出明显的局限性。同时，宏观经济因素（如通货膨胀、利率政策、信用环境等）对金融市场的影响日趋显著，如何有效地将这些宏观经济信号融入资产配置策略中成为了学术界和投资界关注的重要问题。近年来，以机器学习为代表的先进算法，尤其是随机森林（Random Forest）算法，因其强大的非线性处理能力与特征重要性挖掘优势，为动态资产配置提供了新的视角与工具。

* 1. 研究目的与意义

本研究旨在构建并实证验证一种基于随机森林算法的动态资产配置策略，通过深入挖掘宏观经济因子对市场状态的预测能力，实现在不同市场环境下的资产配置优化。具体而言，本研究尝试使用中国市场特有的宏观经济数据（如CPI、国债收益率曲线、信用利差、政策利率等）对未来市场走势进行预测，以此动态调整资产组合，力图提升投资组合的风险调整后收益率。研究结果不仅丰富了资产配置领域的理论研究，也为实际投资者提供了一套科学高效的资产配置框架，具有重要的理论价值和现实意义。

* 1. 报告结构安排

本报告的内容框架安排如下：本报告共六章，结构遵循“问题提出—理论方法—数据与模型—策略实现—实证检验—结论展望”的逻辑递进：第1章引言阐述宏观驱动的动态资产配置背景、研究意义与全文框架；第2章回顾经济周期-资产收益的经典理论，并解析随机森林算法机理及其金融优势；第3章说明宏观指标与多资产价格的数据来源、清洗与标准化流程，并给出增长、通胀、政策三大因子及下行风险标签的构造；第4章详述滚动窗口划分、SMOTE-ENN处理、参数调优与模型评估，输出月度下行风险概率信号；第5章基于该信号设计“概率-权重”线性映射，结合调仓阈值与交易费用，在Backtrader上实现完整回测，最后以收益-风险指标、阶段净值曲线和信号-收益关系评估策略优劣；第6章总结经济直觉、反思局限并提出后续改进方向。

1. 文献综述与理论基础
   1. 文献综述

近年来，伴随机器学习技术的快速发展，特别是随机森林（Random Forest）算法的广泛应用，为动态资产配置提供了新的解决方案（Blitz et al.，2023）。Shu等人（2024）提出一种“资产特有的经济政权预测”框架，结合无监督（jump model）与监督学习（梯度提升决策树），分别对每类资产（如股票、债券、商品）进行周期判断，实现跨资产多阶段动态权重优化。动态资产配置策略旨在通过及时识别并应对宏观经济周期变化，以实现资产组合风险与收益的动态平衡。传统的静态资产配置策略（例如经典的60/40投资组合）虽然长期表现稳健，但在高通胀、政策紧缩或市场剧烈波动的特殊宏观环境下表现欠佳（Mueller-Glissmann & Ferrario，2024）。

Mueller‑Glissmann&Ferrario（2024）利用随机森林模型识别美国宏观周期下的通胀、增长、政策及风险偏好状态，实现对不同资产组合未来尾部风险与反弹概率的预测，并据此动态配置资产。实证表明其方法在高通胀环境下改善了风险调整后收益 。

Dada等（2025）探索了人工智能自适应配置方式，尤其在波动市场以深度学习和强化学习模型预测资产权重调整，实现动态性优化。“增强型资产管理”（2024）评估了AI/ML在资产配置中的透明性与解释能力，强调解释性对于投资者信任与合规管理的重要性。Tetereva等（2024）提出“资产配置森林”（Asset Allocation Forest），改进随机森林对不同资产类（股债配）的非线性处理能力，提升波动性和协方差预测。

* 1. 理论基础

动态资产配置策略的理论基础主要源于经济周期理论和现代投资组合理论（Modern Portfolio Theory，MPT）。经济周期理论认为，经济会经历复苏、扩张、衰退和萧条几个阶段，而不同资产在这些阶段的表现存在明显差异（Ang & Bekaert，2004；Sheikh & Sun，2012；Nystrup et al.，2017）。例如，经济复苏阶段往往伴随股票资产的强劲表现，而经济衰退阶段债券资产通常表现较优（Nystrup et al，2015）。因此，通过及时识别经济周期的位置并据此调整资产配置，可以有效提高资产组合的风险调整收益。

现代投资组合理论则强调通过资产多样化来实现风险与收益的最优权衡（Markowitz，1952）。动态资产配置策略在MPT基础上进一步拓展，通过周期敏感的资产配置，更好地利用周期性资产收益特征，以期在不同市场环境下获得更稳定的风险调整后收益。

随机森林算法作为一种强大的监督式机器学习方法，其优势在于能够处理复杂的非线性数据，捕捉多维数据特征之间的交互作用，并具有较高的模型解释性（Breiman，2001；Géron，2022）。与传统线性模型不同，随机森林算法不需要假设变量之间的线性关系，且具有更强的泛化能力，这使其特别适用于经济与金融领域复杂的数据环境（Yazdani，2020；Zhou et al.，2018）。

Mueller‑Glissmann & Ferrario方法中，使用监督学习明确分类未来尾部风险或上涨概率，这种可控性相比无监督模型（如HMM、聚类）更具实用价值。在资产配置领域，随机森林算法的应用主要通过构建预测模型，来识别和预测宏观经济周期的不同状态，并据此动态调整资产配置策略，以提高投资组合的整体表现（Mueller-Glissmann & Ferrario，2024）。AI 驱动资产配置强调post‑hoc分析和Shapley值解释，帮助投资者理解模型决策逻辑和保护策略透明度 。

1. 数据来源及预处理
   1. 数据来源及说明
      1. 宏观经济数据

本研究使用的宏观经济数据均来源于Wind数据库，涵盖了反映中国经济运行的多个核心指标，包括通货膨胀指标（CPI、原油期货合约价格）、货币与利率指标（AAA级公司债利率、10年期国债利率、2年期国债利率、3月期国债收益率、政策利率）、经济增长指标（制造业PMI、GDP现价）。数据频率为月度，时间跨度自1990年至2024年，覆盖中国及全球经济周期的重要阶段，为动态资产配置策略提供了坚实的宏观经济基础。

* + 1. 资产价格数据

资产价格数据亦来源于Wind数据库，主要包括沪深300指数（HS300）、中债固定利率国债全价指数（IEF）、黄金期货合约结算价（GLD）和原油期货收盘价（WTI），分别代表股票、债券、现金类资产和大宗商品。资产价格数据频率与宏观经济数据保持一致，均为月度数据，以便进行有效的对比分析。

|  |
| --- |
| 图1：归一化资产价格走势 |
|  |
| 资料来源：天风证券研究所 |

* 1. 数据预处理

在数据清洗过程中，首先对数据进行了初步的质量检查，识别并处理了异常值和缺失值。对于异常值，本研究将其视为缺失值；对于缺失值，则采用线性插值法进行填补，以保证数据的连续性和完整性。最后，在合成特征因子之前，对所涉及的宏观经济变量取Z-score，以保证不同量纲的经济变量之间的可比性。经过初步清洗后的宏观经济数据与资产价格数据，按照日期进行了合并，生成了综合数据集。

* 1. 特征工程

在本研究中，我们基于宏观经济理论与机器学习建模实践，构建了多维度的输入特征变量矩阵，通过计算经济增长分数、通胀分数以及政策分数，分别对应经济周期的不同维度，进而全面刻画影响资产收益的宏观因素。整体而言，特征被分为三类：增长因子、通胀因子和政策因子。我们对涉及宏观经济变量绘制了如下相关矩阵热力图。

|  |
| --- |
| 图2：宏观经济数据相关矩阵（2010-2025） |
|  |
| 资料来源：天风证券研究所 |

我们注意到， CPI与WTI的相关性反而为轻微负值，提示能源价格对我国消费物价的传导受到汇率、税费及补贴等多重因素削弱，其成本推动效应并不线性。WTI与其他宏观变量整体相关性较弱且偏负，表明原油价格更多受外部供需和地缘冲击驱动，对国内宏观景气的同步性有限。

此外，我们发现样本期内GDP的上行伴随着制造业繁荣，但伴随着货币政策的边际收紧以及轻微的紧缩，反映出近年来中国经济“增长修复但价格低迷”的结构性特征。PMI作为经济活动的同步或领先指标，表现稳定。同时，经济扩张时长短期国债利差扩大，并与物价短期走势呈反向关系。而政策利率偏差值与长短期国债利差呈中等正相关，反映货币当局在利差扩大时倾向于边际收紧，以防范过热风险。总体而言，样本期呈现“需求复苏-价格温和、政策回归常态”的结构性特征，PMI作为同步／领先指标表现最为稳定。

首先，我们使用中国GDP同比增长以及中国制造业PMI合成所需增长因子。由于PMI更多的作为一种跟随变量，我们将分配更多的权重给GDP同比增长。

|  |  |  |
| --- | --- | --- |
| 图3：中国GDP同比增长（%，1992-2025） |  | 图4：中国制造业PMI（2008-2025） |
|  |  |  |
| 资料来源：天风证券研究所 |  | 资料来源：天风证券研究所 |

其次，我们使用居民消费价格指数（CPI）同比增长与原油价格（WTI）合成通胀因子，前者代表内生性通胀，而后者则代表输入型通胀。在通胀因子中，CPI占主导地位，而WTI则作为补充。

|  |  |  |
| --- | --- | --- |
| 图5：中国CPI指数同比增长（%，1990-2025） |  | 图6：原油WTI价格（2018-2025） |
|  |  | 图表, 折线图  AI 生成的内容可能不正确。 |
| 资料来源：天风证券研究所 |  | 资料来源：天风证券研究所 |

最后，我们使用国债收益率曲线斜率（10年期国债收益率与短端利率之差）与财政扩张（政策利率偏差值）合成政策取向因子，前者代表货币政策取向，而后者则代表财政政策取向。在政策取向因子中，我们依照学术惯例以及货币政策对中国金融市场的作用机制，为国债收益率曲线斜率分配更高的权重。

|  |  |  |
| --- | --- | --- |
| 图7：中国长短期国债收益利差（%，2000-2025） |  | 图8：中国政策利率偏差值（%，2004-2025） |
|  |  |  |
| 资料来源：天风证券研究所 |  | 资料来源：天风证券研究所 |

上述指标经标准化（z-score）处理后构成以下三类宏观特征因子：

Growth\_Score：基于GDP增速与制造业PMI衡量经济增长强度；Inflation\_Score：基于CPI与WTI衡量通胀压力；Policy\_Score：基于财政扩张与长短期国债利差衡量政策取向。通过特征因子的分门别类与合理整合，不仅有效降低了模型的维度，还进一步提升了模型的信息性与对宏观经济走势的预测精度。

1. 单目标随机森林模型
   1. 标签定义与数据标记

针对单目标随机森林策略：我们从原始数据集中提取并滚动标准化宏观经济变量，合成三大宏观因子，并对资产序列取对数收益并构建60/40加权组合。并根据该组合未来一年的累计收益来进行未来回撤是否出现的标记。

|  |
| --- |
| 图9：单目标随机森林模型数标签生成流程图 |
|  |
| 资料来源：天风证券研究所 |

首先，在24期滚动窗口内对每项宏观经济指标计算Z-Score，以消除量纲差异和时序漂移；随后根据相关性与经济含义，将这些标准化指标加权合成为增长、通胀、政策取向三大宏观因子。同时，对沪深300指数与国债的对数收益率进行处理后，以60%沪深300指数/40%国债的静态权重构建基准组合；在每个时点向前滚动12个月，计算该组合的累计收益，若区间累计收益为负，则标记为1，否则为0。最终得到一张包含三大宏观因子、组合收益特征及单一风险标签的训练集，用以训练单目标随机森林模型专注预测未来回撤风险。

* + 1. 滚动训练单目标随机森林分类器

我们采用滚动窗口重训练机制构建预测模型，以充分使用有限时间长度的月度序列。考虑宏观预测任务中数据稀疏与时间序列性质的特点，设定初始训练窗口长度为约5年（60期），从样本开始处依次向前滚动，每次前移1期，逐期进行预测。每轮模型只用前t期进行训练，对t+1期预测，避免数据泄露；模型从第k+1个样本开始运行（k为SMOTE的邻居数）确保小样本类别可用。

本模型采用RandomForestClassifier作为分类模型。同时，我们采用滚动训练过程中固定超参数组合与每轮评估预测AUC与分类精度的方法来主动调参。对于不均衡样本，我们结合SMOTE与Edited Nearest Neighbours的SMOTEENN方法，在每轮训练前判断是否具备过采样条件（小类样本数>k），若不满足则跳过采样。对于缺失值，我们使用中位数填补方法，在事先替换掉所有inf/-inf为空缺值后统一处理。

* + 1. 模型训练效果

图 10 的 ROC 曲线显示，单目标随机森林识别未来回撤风险的能力处于中等偏上水平。图11将模型输出的风险概率与实际12个月累计收益对散点，红色对数拟合曲线呈显著负斜率，表明概率越高，对应收益越低（甚至转负），验证模型概率不仅能区分标签，也具备一定强度的连续信号，可为后续权重映射或动态风控提供量化依据。

|  |  |  |
| --- | --- | --- |
| 图10：滚动窗口ROC曲线与AUC |  | 图11：模型预测效果 |
|  |  |  |
| 资料来源：天风证券研究所 |  | 资料来源：天风证券研究所 |

1. 双目标随机森林模型
   1. 标签定义与数据标记

在双目标随机森林框架中，数据处理和标签构造分两条并行链路完成：首先是宏观因子链——先从原始月度数据中提取十余项核心宏观经济变量（如通胀、产出缺口、利差等），对每个变量按24期滚动窗口计算Z-Score；随后将同类指标加权合成为增长、通胀、政策取向三大宏观因子，为模型提供跨周期、去量纲的经济状态特征。其次是资产收益链——同步读取资产序列的月度收盘价，先取对数收益率，再按照60% 股票/40%国债的静态权重叠加生成基准组合收益；在此基础上累计两段关键标签：1）未来6个月累计收益，收益标记为1，负值记为0，用来捕捉上涨信号；2）未来12个月最大回撤，若区间最大回撤小于-10% 记为1，否则 0，用来衡量潜在风险。最终，每条时间记录同时带有宏观特征矩阵+资产收益特征+2个监督标签，使随机森林能在同一套特征空间下并行学习“获取正收益”与“规避深度回撤”两种信号。

|  |
| --- |
| 图12：双目标随机森林模型数标签生成流程图 |
|  |
| 资料来源：天风证券研究所 |

具体做法如下：首先计算HS300与IEF的月度对数收益；其次按照60%/40%权重合成组合月收益；随后采用6期滚动窗口对组合收益率进行累积，并整体向前平移6期以得到未来6个月的累计收益，同时采用12期滚动窗口得到未来12个月的最大回撤；最后，通过未来6个月累计收益得到收益信号以及通过未来12个月最大回撤得到回撤信号。该标签体系不仅仅能够以投资组合视角衡量跨期下行风险，更能进一步捕捉跨期上涨机会。

* + 1. 滚动训练双目标随机森林分类器

该模型的滚动训练方法与前文所述单目标随机森林模型完全一致，并采用RandomForestClassifier作为分类模型，直接固定训练参数，分别训练出两个针对回撤信号以及上涨信号的随机森林分类器。

两个模型的滚动窗口在滚动窗口检验中，风险模型表现稍优：AUC约0.74、整体准确率约69%，对“未来12个月高回撤”事件召回率77%（漏报少），但精确率仅51%（假阳较多）。收益模型AUC约0.69、同样准确率约69%，对“未来6个月累计正收益”事件的识别精确率76%、召回率64%，表现更均衡。以下左图显示风险模型的ROC曲线：曲线在低假阳率区段快速上升，说明模型对“高回撤”事件具有较强判别力。右图为收益模型，ROC曲线比基准线明显向左上偏移，但与风险模型相比略逊一筹。两条曲线均明显优于45° 随机基线，表明两模型都能提供有用的方向性信号，其中风险模型区分度更高。

|  |  |  |
| --- | --- | --- |
| 图13：风险模型滚动窗口ROC曲线与AUC |  | 图14：收益模型滚动窗口ROC曲线与AUC |
|  |  |  |
| 资料来源：天风证券研究所 |  | 资料来源：天风证券研究所 |

1. 策略实现与回测
   1. 回测资产序列选择

对于回测资产序列，我们使用wind api统一获取获取最新的ETF资产行情数据，并封装为Backtrader可读的数据对象，一次性生成股票类ETF：上证指数ETF（510210.SH）、国债类ETF：国债ETF（511010.SH）、货币类ETF：银华日利ETF（511880.SH）、大宗商品类ETF：大宗商品ETF（510170.SH）、黄金ETF：黄金ETF（518880.SH）五条资产数据流。最后，清洗整理后的多资产价格序列被规范化地送入回测引擎，为后续的动态资产配置策略提供了完整、可追溯的输入数据集。

同时，我们将2013-2025年五类ETF资产的归一化价格与重大宏观事件对齐后绘制于同一张曲线上：1）市场剧震阶段——2015年“股灾”自6月顶部至8月急挫逾 40%，沪深相关的大宗商品指数同步翻倍后归零；2018 年7–12月中美贸易战关税轮番落地，股票、商品再次共振下探，而利率债与货币基金曲线近乎水平，凸显避险属性。2）供给与通胀冲击——2016年一季度的供给侧改革与 2022年2–7月俄乌冲突，两次皆带动大宗商品与黄金同向飙升：GSCI指数阶段涨幅分别约 60%与50%，黄金则滞后确认避险。3）流动性与信用事件——2020年2–4月疫情全球蔓延引发V形抛售，黄金于3月末夺回失地并创新高；2023年3月SVB/CS风波导致黄金三周内跳涨10%，而国债曲线稳步抬升提供底仓稳定。（4）利率预期与主题泡沫——2024年一季度“美联储降息交易”推动股票与商品率先反弹，黄金横盘整理；2025 年二季度AI主题估值回调，股指回撤约15%，但黄金与债券几乎无波动，再次验证跨资产分散效果。总体来看，股票-商品承担进攻、国债-货币充当防守、黄金在尾部风险和通胀冲击中提供保险，多资产组合的动态对冲逻辑由此清晰可见。

|  |
| --- |
| 图15：回测资产序列归一化价格（2013-2025） |
| 图表, 折线图, 直方图  AI 生成的内容可能不正确。 |
| 资料来源：天风证券研究所 |

* 1. 单目标随机森林动态配置策略
     1. 资产权重动态配置决策规则

我们的动态资产配置逻辑不仅仅依赖0/1标签进行“进攻/防御”两档切换，而是把随机森林模型输出的预测概率（未来十二个月回撤概率）P直接映射为五大资产的连续权重。

表1：五大资产的连续权重映射规则（P为风险概率）

|  |  |
| --- | --- |
| 资产 | 权重映射规则 |
| 股票（Stock） | 0.6-0.4P |
| 债券（Bond） | 0.4+0.1P |
| 现金（Cash） | 0.1P |
| 大宗商品（Gsci） | 0.1P |
| 黄金（Gold） | 0.1P |

资料来源：天风证券研究所

由上表看出，资产组合中的股票权重将随风险概率增长从60%线性降至20%，而债券权重则在会高风险期升至最高50%。同时，我们使用现金和大宗商品分别引入流动性缓冲与输入型通胀对冲。而黄金则用剩余权重自动补足，使总和恒为1。

本策略按月再平衡：仅在每月首个交易日检查随机森林给出的最新回撤概率P，若与上月差幅大于等于1百分点才调仓；触发调仓时先清空所有头寸，再按新权重等值买入五类ETF；启动日默认持60/40股债组合，其余交易日仅记录净值。

* + 1. 回测参数框架设计与交易逻辑实现

我们的回测部分采用Backtrader框架，起始资金设为1000000元，回测区间覆盖2014-01-01至2025-07-01；基准资产文件分别对应沪深300股指ETF、中长期国债全价指数ETF、货币ETF、大宗商品ETF与黄金ETF的日频数据中的收盘价作为回测数据。费用假设为零佣金零滑点，日无风险利率固定为1%。

交易逻辑由随机森林模型的风险概率驱动，当月度首个交易日最新计算的风险概率与上月值绝对差≥1%时才触发调仓，先清仓、再按上述权重等值建仓，以减少换手占优。模型训练、预测与调仓过程完全嵌入 Backtrader的MacroDynamicProbStrategy类，回测结束自动保存每日净值并通过该平台自带分析器类函数输出年化收益、最大回撤和夏普比率等绩效指标，为后续性能评估提供数据基础。

* + 1. 策略整体绩效评估

在完成回测之后，我们进一步汇总了基于单目标随机森林模型的动态资产配置策略的关键绩效指标如下。

表2：单目标随机森林模型动态资产配置策略关键绩效指标（按年）

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 年化收益(%) | 年化波动(%) | 最大回撤(%) | Sharpe | Calmar | Sortino |
| 2014 | 31.58 | 13.05 | -4.42 | 2.25 | 7.15 | 3.37 |
| 2015 | 4.25 | 22.31 | -23.73 | 0.45 | 0.18 | 0.59 |
| 2016 | -1.01 | 14.29 | -10.33 | -0.11 | -0.10 | -0.14 |
| 2017 | 5.55 | 7.19 | -3.50 | 0.82 | 1.59 | 1.28 |
| 2018 | -7.66 | 10.74 | -11.01 | -0.69 | -0.70 | -0.89 |
| 2019 | 16.75 | 9.15 | -6.14 | 1.59 | 2.73 | 2.37 |
| 2020 | 12.87 | 9.20 | -5.81 | 1.44 | 2.22 | 2.03 |
| 2021 | 6.93 | 5.87 | -4.82 | 1.28 | 1.44 | 1.91 |
| 2022 | -2.83 | 7.27 | -6.67 | -0.38 | -0.43 | -0.50 |
| 2023 | 0.99 | 4.97 | -5.14 | 0.32 | 0.19 | 0.53 |
| 2024 | 9.26 | 10.27 | -5.92 | 0.93 | 1.57 | 1.29 |
| 2025 | 4.24 | 7.27 | -4.30 | 0.94 | 0.99 | 0.83 |
| 汇总 | 6.84 | 11.24 | -25.32 | 0.64 | 0.27 | 0.81 |

资料来源：天风证券研究所

纵观2014-2025年的绩效序列，策略呈现先扬后抑、周期性波动的特征：2014年录得31.58%的年化收益、仅-4.42%的最大回撤，Sharpe和Calmar分别为2.25、7.15，显示在单边牛市中捕捉趋势且控制风险的能力极强；随后在2015年股灾（收益4.25%，最大回撤-22.31%，Calmar为0.18）与2018年中美贸易战、2022年俄乌冲突（整体收益分别回撤-7.66%与-2.83%，Sharpe、Sortino均为负）三次外部冲击中暴露出策略防守不足的问题，为其长期表现的主要拖累。2016年略亏、2017-2018年与2021-2025年上半年大多处于0.99%-9.26%的温和收益区间，波动在4.97%-10.27%，Sharpe为-0.38-1.28，体现“低波保值”属性；其中2019-2020年，该策略在风险偏好回升阶段再次贡献16.75%和12.87%的高收益，波动仅9%左右，Sharpe大于1.4。最终12年复合年化收益6.84%，年化波动11.24%，最大回撤-25.32%，Sharpe为0.64、Sortino为0.81，Calmar为0.27；总体说明策略在顺风行情具备捕捉Alpha的能力，但极端市场期间缺乏及时减仓机制，单次深度回撤显著拉低了长期“收益/风险”效率，未来若能引入动态防守或尾部风险对冲，Calmar与稳健度仍有较大提升空间。

* 1. 单目标随机森林动态配置策略
     1. 资产权重动态配置决策规则

我们的动态资产配置逻辑不仅仅依赖0/1标签进行“进攻/防御”两档切换，而是把随机森林模型输出的预测概率（未来6个月累计正收益的概率R以及未来十二个月出现重大回撤概率K）共同直接映射为五大资产的连续权重。

表3：五大资产的连续权重映射规则（R为收益模型概率，K为回撤模型概率）

|  |  |
| --- | --- |
| 资产 | 权重映射规则 |
| 股票(Stock) | max（0，min（0.6，0.3+0.3R−0.3K）） |
| 国债(Bond) | Max（0，min（0.6，0.5−0.2R＋0.2K）） |
| 现金(Cash) | 0.1K |
| 大宗商品(Gsci) | 0.1R |
| 黄金(Gold) | Max（0，1−其他资产权重之和） |

资料来源：天风证券研究所

由上表映射公式可知，股票的基准权重为30%，会随着收益概率R上升而增加、随风险概率K上升而减少，在0–60%区间动态浮动；国债以50%为起点，受R负向、K正向调节，最高可抬升至60%，在高风险阶段承担主要防守作用。现金权重按0.1K线性增加，为极端情形提供流动性；大宗商品则按0.1R递增，用作进攻及输入型通胀对冲。组合剩余部分自动配置到黄金，使五类资产权重始终合计为1。

本策略按月再平衡：每到当月首个交易日先用双随机森林模型算出“上涨概率R”和“回撤概率K”。若两者相对上月的变化均不足5个百分点，则维持原仓位；否则重算权重并再平衡。调仓时先清空旧仓后按新权重等值买入五只ETF。策略启动日默认持有60/40股债基准组合，其余交易日仅记录净值，回测结束输出累计收益曲线。

* + 1. 回测参数框架设计与交易逻辑实现

我们的回测部分采用Backtrader框架，起始资金设为1000000元，回测区间覆盖2014-01-01至2025-07-01；基准资产文件分别对应沪深300股指ETF、中长期国债全价指数ETF、货币ETF、大宗商品ETF与黄金ETF的日频数据中的收盘价作为回测数据。费用假设为零佣金零滑点，日无风险利率固定为1%。

交易逻辑由随机森林“双概率”驱动：在每月首个交易日，策略先用宏观特征喂入回撤模型得出风险概率 K，用收益模型得出上涨概率 R；只有当 K 或 R 相对上月任一变动≥5%才触发再平衡。触发时先全盘平仓，再按连续映射公式补足权重等值建仓，以在不显著增加换手的前提下动态切换进攻/防守。模型加载、概率计算与调仓流程全部封装在 Backtrader的DualMacroStrategy类中；回测期间每日记录账户价值，结束后保存净值曲线，再通过Backtrader内置分析器计算年化收益、最大回撤、Sharpe等策略绩效，为后续评估与可视化提供完整数据。

* + 1. 策略整体绩效评估

在完成回测之后，我们进一步汇总了基于双目标随机森林模型的动态资产配置策略的关键绩效指标如下。

表4：双目标随机森林模型动态资产配置策略关键绩效指标（按年）

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 年化收益(%) | 年化波动(%) | 最大回撤(%) | Sharpe | Calmar | Sortino |
| 2014 | 24.36 | 10.66 | -4.12 | 2.17 | 5.91 | 3.2 |
| 2015 | 1.96 | 12.86 | -14.87 | 0.42 | 0.13 | 0.54 |
| 2016 | 6.76 | 12.17 | -5.4 | 0.55 | 1.25 | 0.78 |
| 2017 | 6.11 | 6.2 | -3.57 | 1.03 | 1.71 | 1.58 |
| 2018 | -5.25 | 9.19 | -7.3 | -0.54 | -0.72 | -0.66 |
| 2019 | 19.17 | 8.27 | -4.95 | 1.99 | 3.88 | 3.03 |
| 2020 | 12.71 | 10.26 | -8.1 | 1.3 | 1.57 | 1.75 |
| 2021 | 2.29 | 3.56 | -3.6 | 0.83 | 0.64 | 1.05 |
| 2022 | -0.4 | 5.24 | -5.31 | -0.08 | -0.07 | -0.12 |
| 2023 | 2.07 | 3.81 | -4.81 | 0.67 | 0.43 | 0.96 |
| 2024 | 8.4 | 10.02 | -5.63 | 0.88 | 1.49 | 1.2 |
| 2025 | 3.45 | 8.32 | -5.18 | 0.64 | 0.67 | 0.56 |
| 汇总 | 7.05 | 8.91 | -14.87 | 0.81 | 0.47 | 1.04 |

资料来源：天风证券研究所

纵观2014-2025上半年绩效表现，该策略呈“高光-回撤”交替特征：2014年年化收益24.4%、最大回撤仅-4.1%，Sharpe为2.17、Calmar为5.91，凸显在单边多头中的高效进攻；2019年的表现同样亮眼，收益为19.2%、波动仅8.3%，Sharpe接近2，风险收益比优越。相对地，2015年“股灾”阶段策略收益缩至2%且最大回撤放大到-14.9%，Calmar跌至0.13；中美贸易战爆发的2018年与俄乌冲突爆发的2022年，该策略均录得负收益（-5.3%、-0.4%）且Sharpe、Sortino转负，显示策略在外部冲击下仍缺乏足够防守，但其回撤控制明显优于单目标随机森林策略。其余年份（2016-2017年、2020-2021年及2023-2025年）收益大多介于2%-13%，波动4%-12%，Sharpe为0.5-1.3，体现策略低波稳健的特点。整体12年复合年化收益7.05%，年化波动8.91%，最大回撤为-14.9%，Sharpe为0.81、Sortino为1.04、Calmar为0.47：相比单目标策略，尾部损失显著降低，但在高风险年份仍有回撤偏大的问题；若进一步引入动态减仓或对冲机制，预计可继续抬升Calmar和长期稳健度。

* 1. 基准策略及多策略对比分析
     1. 基准策略组合配置决策规则

策略遵循“先买后抱”单一决策规则：在回测启动后的首根K线开盘时，按预设权重——60%股票ETF、40%国债ETF，其余资产权重为0%——次性买入，买入量用账户实时市值×权重÷当日开盘价计算并向下取整；若某资产开盘价缺失或为0则跳过。建仓后全过程不再调仓，直至回测结束由Backtrader自动平仓。每日收盘后记录组合市值并计算累计收益，最终输出完整净值曲线用于评估。

* + 1. 策略整体绩效评估

在完成回测之后，我们进一步汇总了基准策略的关键绩效指标如下。

表5：基准策略关键绩效指标（按年）

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 年化收益(%) | 年化波动(%) | 最大回撤(%) | Sharpe | Calmar | Sortino |
| 2014 | 36.23 | 16.49 | -5.35 | 2.03 | 6.77 | 2.93 |
| 2015 | 1.74 | 34.12 | -34.3 | 0.34 | 0.05 | 0.44 |
| 2016 | -4.52 | 19.94 | -15.6 | -0.24 | -0.29 | -0.29 |
| 2017 | 8.05 | 12.4 | -5.4 | 0.71 | 1.49 | 1.06 |
| 2018 | -14.74 | 16.68 | -19.19 | -0.88 | -0.77 | -1.12 |
| 2019 | 20.31 | 13.27 | -9.07 | 1.34 | 2.24 | 1.95 |
| 2020 | 16.54 | 15.2 | -9.34 | 1.16 | 1.77 | 1.52 |
| 2021 | 4.07 | 10.41 | -7.14 | 0.51 | 0.57 | 0.77 |
| 2022 | -9.69 | 12.47 | -14.27 | -0.79 | -0.68 | -1.04 |
| 2023 | -0.35 | 8.97 | -9.66 | 0.09 | -0.04 | 0.16 |
| 2024 | 9.46 | 17.57 | -10.92 | 0.6 | 0.87 | 0.84 |
| 2025 | 6.7 | 12.2 | -7.3 | 0.81 | 0.92 | 0.7 |
| 汇总 | 5.77 | 17.2 | -37.32 | 0.41 | 0.15 | 0.51 |

资料来源：天风证券研究所

纵观2014-2025上半年绩效表现：2014年策略以36.23%的高收益、仅-5.35%的最大回撤和2.03的Sharpe起步，显示强劲进攻能力，但随后风险管理短板暴露：2015年股灾和2018年中美贸易战爆发导致该策略最大回撤分别扩大至-34.3%与-19.2%，年化波动飙至34.12%和16.68%，Sharpe均跌至0.3以下；2022年俄乌冲突爆发后，该策略亦录得-9.69 %的亏损。该策略亮点主要集中在2019-2020年：收益分别为20.31%、16.54%，Sharpe超过1，Calmar超1.7，证明该策略在多头行情下仍具捕捉alpha的能力。其余年份基本维持4-9%区间的温和收益。整体12年，该策略复合年化收益仅5.77%，年化波动高达17.2%，最大回撤-37.32%，Sharpe为0.41、Calmar为0.15；说明策略虽能在牛市阶段大幅做多，但缺少有效的减仓或对冲机制，导致极端行情下巨幅回撤严重稀释长期风险调整回报。

* + 1. 多策略对比分析

综合三种策略的回测结果，我们得到在不同市场环境下的归一化策略累计净值对比如下图所示。

|  |
| --- |
| 图16：不同市场环境下多策略归一化收益对比（2014-2025） |
|  |
| 资料来源：天风证券研究所 |

图16比较了基准60股/40债组合（橙色虚线）与两套机器学习策略的归一化累计净值曲线：单目标随机森林（蓝色实线）和双目标随机森林（绿色点划线）。整体来看，两种随机森林策略均显著跑赢基准，且双目标曲线始终位于最上方：在2014-2015年牛市区（绿色底纹）三者同步上冲，但2015年股灾后基准净值急挫至1.2附近，而双目标策略仅回落到约1.4，展现更强抗跌性；震荡区（浅黄色底纹）内，双目标曲线保持缓坡上行，单目标稍次，基准则横向拉锯。进入2019-2020年新一轮多头，双目标策略升穿1.8，并在2022-2023年熊市区（红色底纹）中回撤受控、快速刷新高点，而基准深幅回撤后恢复乏力。至 2025年末，双目标净值逼近2.2，单目标约2.0，基准仅1.8左右；多次熊市阴影下的“V”形走势说明机器学习策略能较好削减尾部损失并加速再创新高，尤其双目标模型在风险空窗期捕捉进攻机会、更充分利用防守资产，从而取得最优的收益-回撤平衡。

1. 研究结论
   1. 策略表现讨论与局限性分析
      1. 策略效果的经济直觉与解释

本报告所构建策略的本质是在宏观视角下动态搬移风险：

单目标随机森林模型策略仅使用随机森林将通胀、增长、政策等多维宏观变量映射为一条 “下行风险概率K”，并据此连续调整股票、债券、现金、大宗和黄金权重——K高则压低股票、抬升债券及避险资产，K低则恢复权益敞口；以月频滚动，且仅当K较上月变动超过阈值时才调仓，从而减少换手与滑点。双目标随机森林模型策略在此基础上再引入 “正收益概率R”，权重同时随R正向、随K反向平滑变化，使组合能在宏观回暖时更及时加仓进攻资产，在风险抬升时更彻底收缩。

* + 1. 策略局限性与改进方向

尽管策略整体表现优异，但仍存在以下局限：

1）是信息滞后，月度宏观数据发布滞后2-4周，需通过高频替代指标和now-casting方法弥补；2）是随机森林忽视时序依赖，面对结构突变适应性不足，可引入LSTM/Transformer或递减加权树提升响应；3）是SMOTE-ENN可能导致概率校准偏差，应采用Platt/Isotonic校准并动态调阈；4）是权重映射系数固定，可用贝叶斯优化或强化学习实现自适应；5）是回测摩擦简化，实盘应加入深度-冲击模型并让调仓阈值与成本联动；6）是资产池仅含股债现金商品黄金，后续可扩展至可转债、REITs、外汇等提高分散度；7）是过度依赖宏观视角，需融合资金流、期权波动率与情绪信号构建多模态框架。

* 1. 研究结论及总结

本研究以增长、通胀与流动性等宏观数据为核心信息源，构建了单目标（仅预测回撤概率K）与双目标（同时预测回撤概率K与上涨概率R）两套随机森林动态资产配置框架，并将概率连续映射为股票、国债、现金、大宗商品、黄金五类ETF的权重；月度滚动、阈值调仓设计在降低换手的同时使组合能够随宏观状态在“进攻——防守”之间平滑切换。2014-2025年回测结果显示：单目标策略实现年化收益6.82%、波动率11.24%、Sharpe0.64，已明显优于静态60/40组合（5.77%、17.20%、0.41）；双目标策略进一步提升至年化收益7.05%、波动率8.91%，最大回撤从 −37.3%压缩至−14.9%，Sharpe升至0.81，凸显将R与K联合纳入权重决策可在扩张期充分捕捉风险溢价、在冲击期有效压降尾部损失。与传统静态配置相比，本框架在多轮牛熊切换中稳定提升风险调整后收益，验证了“宏观状态识别 → 概率量化 → 连续权重控制”闭环的可行性和经济直觉一致性；同时也揭示，信息时滞、模型时序假设与交易摩擦仍是影响实盘表现的关键瓶颈，未来可通过高频替代数据、序列模型融合、自适应权重函数及更多元资产扩展来进一步优化组合韧性与收益质量。