

让情绪“有结构”：大模型如何挖掘研报新价值

——量化研究系列报告之二十三

报告日期：2025-08-11

分析师：骆昱杉

执业证书号：S0010522110001

邮箱：luoyushan@hazq.com

分析师：严佳炜

执业证书号：S0010520070001

邮箱：yanjw@hazq.com

相关报告

- 《临界相变：探寻传统因子中的非线性基因——量化研究系列报告之二十二》2025-06-13
- 《ETF 资金流透视：被动化浪潮下行业与个股的演进——量化研究系列报告之二十一》2024-12-25
- 《上市板块差异性视角下的选股策略优化与实战指南——量化研究系列报告之二十》2024-12-23
- 《破解 Alpha 投资困境：因子择时方案再探索》——量化研究系列报告之十九》2024-09-17
- 《企业利润分配策略：短期股东回报与长期价值创造的平衡》——量化研究系列报告之十八》2024-08-14
- 《另辟蹊径：发掘分析师因子中的另类 alpha——量化研究系列报告之十七》2024-06-13

主要观点：

本篇报告依托 DeepSeek 大模型，将情感分析从研报标题拓展至摘要层面，结合摘要结构解析与多维量化方法，并基于结构化情绪信号实现对研报摘要价值的重构。

● 基于 DeepSeek 大模型的研报摘要结构化情感分析

本文依托 DeepSeek 大模型，将情感分析从研报标题拓展至摘要，实现从黑箱评分向结构化、可解释评分的转变。通过拆解摘要结构（业务发展—战略布局、技术创新、渗透率、经营效率；财务表现—盈利质量、稳健性、业绩对比；其他因素—股东行为、政策影响、市场驱动），引入类别识别、情绪倾向、出现顺序及字数占比四维量化分析。

● 结构化情绪视角下研报摘要价值的多维洞察

从实证结果看，分类识别显著提升了研报分析的颗粒度，能多维度揭示企业真实价值。其中，利好情绪的“密度”指标对超额收益预测有效，尤其“业绩超预期”密度、“盈利改善”密度表现突出，相关选股策略显著跑赢市场。情绪信息的“出现顺序”与“篇幅占比”反映强调方式，利好越早出现或篇幅越大，定价能力越强，战略、创新及政策类信号尤为明显，但基本面信号得效力较弱。

● 研报摘要因子的构建与多维绩效评估

综合情绪类别、篇幅占比与情绪密度等维度，构建四类研报评分因子（简单加权、篇幅加权、类别重要性加权、篇幅&类别重要性加权），并引入集中度调整与非线性权重进行优化。实证结果显示，研报评分在报告发布后 20 日、40 日具有显著的收益解释力，体现出较强的事件驱动特征。

将四类因子等权合成个股综合评分因子 `score_report_llm`，其整体表现稳健，五分组收益结构严格单调，具备一定选股与收益预测能力，且与传统因子相关性较低。`score_report_llm` 的多头年化收益为 12.8%、相对中证 800 的年化超额 13.5%，自 2020 年以来年胜率为 100%，今年截至 5 月底的月胜率为 100%，超额 10.4%；年度超额最大回撤控制良好，相对等权组合均控制在 4% 以内。

● 风险提示

量化模型基于历史数据，过去的回测业绩不代表未来；量化模型本身存在失效的风险。

正文目录

1 大模型迭代下的文本探索之路	5
2 研报内容的框架拆解与多维信息挖掘	6
2.1 研报摘要内容的结构化解读	6
2.1.1 业务发展：战略优化与核心能力提升	7
2.1.2 财务表现：盈利质量、稳健性与业绩超预期分析	8
2.1.3 其他因素：股东行为、政策影响与市场驱动分析	9
2.2 摘要信息处理与多维度量化分析	9
2.3 模型选择与场景规则化训练	10
2.3.1 模型选择与部署调用	10
2.3.2 任务规则化：为大模型在投资领域“降噪”增效	11
2.3.3 模型调用实践：本地与云端的高效对接	14
3 基于 DEEPSEEK 的分类任务实践与企业价值洞察	15
3.1 DEEPSEEK 在研报分类任务的实践	15
3.2 文本分类准确性的复查与调整	17
3.3 细分类别助力从细节洞察企业真实价值	19
4 结构化情绪视角下的研报摘要价值重构	20
4.1 研报分类与情绪分析全景	20
4.2 利好密度驱动下的收益预测与策略应用	21
4.2.1 短期利好驱动强，超预期与盈利改善信息更具价值	22
4.2.2 从报告到个股：分类利好密度可有效预测未来收益	23
4.2.3 构建基于情绪利好密度构建主动选股策略	24
4.3 研报信息结构是否影响未来走势？	26
4.3.1 顺序：类别信息靠前意味着更强定价效应？	26
4.3.2 篇幅：重点说明 vs 轻描淡写的信号效应	28
5 研报摘要综合因子的构建与评估	28
5.1 报告维度的综合评分算法	28
5.1.1 简单加权及集中度调整	28
5.1.2 篇幅占比加权	29
5.1.3 类别重要性加权	30
5.1.4 篇幅占比&类别重要性加权	32
5.2 个股层面的评分指标构建	33
6 总结	37
风险提示：	38

图表目录

图表 1 CHATGPT 与 DEEPSEEK 模型对比	5
图表 2 分析师撰写的公司研报的典型框架	6
图表 3 研报摘要业绩部分的典型结构	7
图表 4 不同尺寸的 DEEPSEEK-R1 版本介绍、参数量、特点、使用场景和硬件配置	11
图表 5 测试模型的部署方式和调用费用	11
图表 6 研报摘要的利好点分类定义与情绪评分规则	12
图表 7 系统提示	13
图表 8 DEEPSEEK-R1-32B 模型调用的输入输出示例：输出格式不符合规则，不是 JSON 格式	16
图表 9 DEEPSEEK-R1-671B 模型调用的输入输出示例：输出格式符合规则	16
图表 10 DEEPSEEK-V3-671B 模型调用的输入输出示例：输出结果与 R1-671B 相似	17
图表 11 DEEPSEEK-V3-671B 模型基本符合本研究任务	17
图表 12 分类 7（业绩对比预期与同业）的分词逻辑	18
图表 13 分类 7（业绩对比预期与同业）的复查逻辑	18
图表 14 “沙里淘金” 举例	19
图表 15 “去伪存真” 举例	20
图表 16 “拨云见日” 举例	20
图表 17 分年度有效研报数量	20
图表 18 研报涉及类别数统计	20
图表 19 区间内，涉及各类别的报告数量与占比	21
图表 20 分年度，各分类数量占报告数的情况	21
图表 21 各年度的情绪分布（%）	21
图表 22 各分类的情绪分布（%）	21
图表 23 各类别，不同情绪的报告对应的未来 CAR 情况	22
图表 24 个股维度：各分类利好密度回归结果	23
图表 25 个股维度：各分类利好密度回归结果（行业因素）	23
图表 26 盈利改善密度组合的分年度收益情况（月频调仓）	25
图表 27 盈利改善密度组合的分年度收益情况（不定期调仓）	25
图表 28 盈利改善密度组合的超额净值曲线	25
图表 29 业绩超预期密度组合的分年度收益情况（月频调仓）	26
图表 30 业绩超预期密度组合的分年度收益情况（不定期调仓）	26
图表 31 出现顺序与 CAR 的回归结果（情绪值为+1）	27
图表 32 出现顺序与 CAR 的回归结果（情绪值为-1）	27
图表 33 篇幅占比与 CAR 的回归结果（情绪值为+1）	28
图表 34 篇幅占比与 CAR 的回归结果（情绪值为-1）	28
图表 35 score_sum和score_mean_hhi与 CAR 的回归结果	29
图表 36 score_by_len与 CAR 的回归结果	30
图表 37 类别重要性权重的计算方案	31
图表 38 score_by_cat与 CAR 的回归结果	32
图表 39 score_by_LenCat与 CAR 的回归结果	33
图表 40 指数衰减加权的权重计算举例	33
图表 41 指数衰减加权的权重示意图	33

图表 42 个股评分因子及逻辑.....	34
图表 43 <i>score_mean</i> 因子表现.....	34
图表 44 <i>score_by_len</i> 因子表现.....	35
图表 45 <i>score_by_cat</i> 因子表现.....	35
图表 46 <i>score_by_LenCat</i> 因子表现.....	35
图表 47 基于研报摘要分类的合成因子的表现.....	36
图表 48 <i>score_report_llm</i> 的分组年化超额（中证 800）.....	37
图表 49 <i>score_report_llm</i> 的分组超额净值（中证 800）.....	37
图表 50 <i>score_report_llm</i> 因子的多头分年度表现.....	37
图表 51 <i>score_report_llm</i> 因子的相关性.....	37

1 大模型迭代下的文本探索之路

2022年11月，ChatGPT 面世后，我们首次尝试将其应用于情感分析领域，通过分析卖方分析师对公司盈余公告点评标题中的情感态度，取得了初步成果，感受到了大型语言模型在文本分析中的强大潜力。ChatGPT 不仅能够准确且连贯地进行评分，还能精准捕捉核心情感信息，表现明显优于传统模型如 BERT（详见我们 2023 年 8 月 14 日发布的专题报告《ChatGPT 与研报文本情绪的碰撞——量化研究系列报告之十一》）。然而，受限于当时的技术水平与模型调用成本，我们的研究仅聚焦于标题情感分析，尚未延伸至研报摘要部分。

今年，DeepSeek 模型的出现为我们提供了新的契机。该模型相比 ChatGPT，参数规模更小、成本更低、操作更便捷，且依然保持了出色的性能表现，使得大规模自动化文本分析任务成为现实可能。

图表 1 ChatGPT 与 DeepSeek 模型对比

类别	ChatGPT	DeepSeek
模型架构	基于传统 Transformer 架构，所有参数参与任务	采用混合专家 (MoE) 架构和多头潜在注意力 (MLA) 机制，每次请求仅激活部分参数
性能能力	语言理解和生成能力出色，处理复杂语言表达和情感分析表现出色，如在创意写作、文案撰写等任务能生成高质量内容	在数学计算和逻辑推理方面表现出色，如在 AIME2024 数学基准测试中得分高于 OpenAI o1，推理算法经强化学习优化，可进行多步骤推理和思考
训练效率与成本	训练成本较高，如 GPT-4 训练成本高达数亿美元	训练成本较低，如 DeepSeek-V3 训练过程约 55 天，成本 558 万美元
应用场景与灵活性	在教育领域应用广泛，可用于学生查找资料、解答疑问、练习语言等，在创意写作方面能生成高质量创意写作内容	在科研和数学领域优势明显，能快速解决复杂数学问题和进行逻辑推理，在软件开发领域能生成高质量代码片段并提供详细开发指导
数据依赖性	训练数据主要来源于互联网大量文本，对数据质量和多样性高度依赖	数据来源相对集中，主要依赖特定训练数据集
信息准确性与可靠性	训练数据更新频率有限，处理时效性强问题可能无法提供最新答案	推理过程复杂，处理需长逻辑链条任务时，可能出现推理错误或结果不准确
性能稳定性	算法复杂度高，需大量计算资源和时间，限制其实时交互场景应用效果	推理效率高，处理复杂任务能快速给出答案
开源与生态	由 OpenAI 主导，通过 API 和生态合作覆盖全球开发者与企业用户	采用开源策略，吸引众多开发者参与优化和定制，推动技术普及应用
费用模式	分层订阅制，Plus 版每月 20 美元，Pro 版 200 美元，企业版定制报价	完全免费且开源，API 调用成本仅为 ChatGPT 的 1/10，如 R1 模型每百万输出 tokens 约 16 元
数据安全	数据存储于海外服务器，隐私政策复杂，国内用户需 VPN 访问，存在数据跨境风险	支持本地部署，如通过 Ollama 在普通电脑运行，数据可完全本地化存储，适合处理敏感信息

资料来源：华安证券研究所整理

基于此，本研究尝试利用 DeepSeek 将文本分析拓展至卖方分析师研报的摘要部分。一方面，摘要承载了更丰富、全面的信息；另一方面，近年来分析师研报措辞普遍偏积极，特别是标题部分过于简略且不包含客观数据，单纯依赖文字情绪易导致偏差。

传统的诸多研究多聚焦于大模型对摘要进行综合评分，虽取得一定成果，但仍存在若干亟需探讨的问题：其一，评分过程“黑箱化”，模型直接输出总分，难以洞察评分细节；其二，情绪评分易受分析师个人写作风格影响，乐观型分析师偏好正面词汇，部分中立分析师则措辞平和，导致评分可能被局部情绪主导；其三，综合评分缺乏灵活性，通常仅反映分析师对公司整体状况的判断，难以洞察不同经营维度的细节。

因此，本文基于 DeepSeek 大模型，围绕卖方研报摘要展开系统研究，实现“黑

箱评分”向“可解释结构化”评分。我们从分析逻辑与方法论出发，结合模型选择与训练过程，深入解析模型输出的情绪结构与信息维度，并探讨其在投资决策中的实际应用价值。

2 研报内容的框架拆解与多维信息挖掘

在金融研究领域，研报是分析师与投资者进行信息交流的关键桥梁，而摘要浓缩了全文的核心内容。本章将介绍研报分类的整体框架，深入探讨从研报结构分析到情绪分类的全过程，同时详细介绍分类任务中的各类细节，包括业务发展、财务表现、外部因素等方面的分类定义。通过对研报结构的拆解与分类解读，探索其在投资决策中的应用价值和实用性。

2.1 研报摘要内容的结构化解读

分析师撰写的公司报告的摘要通常遵循一个程序化的框架，主要包括以下四个部分：

- **事件：**对核心经营结果的总结，涵盖财务预告、快报或正式报告中的关键财务数据，以及公司的重要事件，如分红公告、股份回购简要和其他重大事件。
- **点评：**分析师对上述“事件”的详细分析和观点，通常包括对公司整体业绩的评价、经营性指标的解读、毛利率和净利率的分析、流动性状况、销售渠道表现、研发进展以及战略布局等内容。此外，还可能涉及股东行为、政策环境、市场资金和热度等其他相关因素。
- **盈利预测：**对公司的评级，以及未来三年的盈利预测、估值等前瞻性分析。
- **风险提示：**对公司可能面临的风险因素进行说明和提示。

图表 2 分析师撰写的公司研报的典型框架

章节	内容
事件/总结	公司发布2023年年报，2023年实现营业收入2.46亿元，同比-6.47%；毛利率19.55%，同比下降5.77pct；归母净利润-1.85亿元，亏损同比+15.97%；扣非归母净利润-1.91亿元，亏损同比+16.37%，非经常性损益主要由政府补助及其他非流动金融资产公允价值变动构成。
点评	加快智慧政务市场拓展，业绩短期承压。 公司多年深耕智慧政务市场，涵盖数字档案和智慧追溯管理平台等多种业务，智慧政务23年营收2.06亿元，同比-8.39%，毛利率为16.96%，同比下降11.23pct。公司在2023年加大了市场拓展力度，但受限于目标客户预算收紧、需求不足，新增项目签约不及预期；同时客户资金紧张导致存量项目验收及回款速度下降，两方面原因导致公司短期内业绩承压。 深度布局智慧交通，需求逐步释放。 公司实控人为交通运输部下属中国交通通信信息中心直属国企交通通信集团，且持有交信北斗（海南）及交信北斗（浙江）51%的股份，智慧交通业务围绕“北斗+”开展，包括北斗自由流里程费、车联网、基建智能化等业务，2023年实现营收0.39亿元，同比+1365.72%；实现毛利0.16亿元，同比+741.33%。公司于重点推进智慧交通业务发展，2023年4月完成了7亿元增发，用于北斗自由流项目及补充流动资金，业务收益初见成效。 低空产业蓬勃发展，陆空一体化系统推进北斗应用。 北斗产业化应用在“十四五”期间进入发展快车道，低空产业迎来快速发展期，国内低空经济市场规模预计超万亿。公司实控人交通运输部通信信息集团致力于推动实施“发展天地一体通信体系、服务综合立体交通信息化”等发展战略，经国务院批准在中国合法提供航空安全通信服务近40年，有望赋能公司发展。公司控股子公司交信北斗（浙江）科技有限公司着力发展新质生产力，结合无人机与北斗卫星系统，开发了一套陆空一体化交通基础设施健康监测系统，通过无人机巡飞及交通基建上的传感器，对桥梁、边坡及健康状况进行实时监测、预警，为清除道路病害及智能养护提供指导；依托低空空域与北斗系统，为智能交通管理提供新解决方案。
盈利预测/投资建议	预计公司2024-2026年的归母净利润分别为-0.95/0.96/2.01亿元，对应25/26年PE分别为33/16倍，给予公司2025E的盈利50倍PE估值，对应目标价19.35元，首次覆盖，给予“买入”评级。
风险提示	盈利预测与估值判断不及预期、业务拓展不及预期

资料来源：慧博投研，华安证券研究所

为了确保研究的准确性和客观性，避免模型过度依赖盈利预测和风险提示部分，进而对分析结果产生干扰，我们对研报摘要的原文进行了初步清洗，剔除了盈利预测和风险提示这两部分内容。

以典型的业绩点评为例，其结构通常包括：先总体描述公司业绩成果，然后对业

绩进行拆分，分析单项构成并陈述分析师的观点，内容涵盖整体盈利情况、公司经营指标、毛利率和净利率、流动性指标、公司销售渠道情况、研发相关情况以及公司的战略布局等。此外，还可能涉及股东行为、政策情况、市场资金和热度等其他相关因素。

图表 3 研报摘要业绩部分的典型结构

研报摘要原文（业绩部分）	说明
业绩大超预期，源于N业务贡献和降本增效公司预计2024上半年实现归属于上市公司股东的净利润为3700到4070万元，同比增长460.15%到516.16%，预计实现扣非净利润3450到3795万元，同比由负转正（去年同期为28.51万元），业绩超市场预期。	整体描述
业绩大增主要是因为： 1) 公司前瞻布局新产品、新赛道、新业务，使得营收实现稳步增长，同时综合毛利率较上年同期回升明显； 2) 海外业务拓展逐步落地，N公司业务实现了从0到1的突破，海外业务收入同比显著增长； 3) 公司对内提高运营效率，降本增效，期间费用同比略有下降，淳中转债完成转股和强赎，减少了债券利息费用。	业绩分析

资料来源：慧博投研，华安证券研究所整理

在综合分析多家卖方、多个类型的研报摘要后，我们将研报摘要中涉及的利好类别分为以下三大类、十个子类：

- 业务发展：战略布局、技术创新、渗透率提升、经营效率提升。
- 财务表现：盈利改善、财务稳健性提升、业绩对比预期与同业。
- 其他因素：股东行为、政策驱动、市场驱动。

2.1.1 业务发展：战略优化与核心能力提升

“业务发展”类聚焦于公司优化举措，不一定已经体现在了披露业绩中，包括战略布局、技术创新、渗透率、经营效率。

战略布局：基于外部环境研判与内部资源评估，企业主动进行或者制定的跨周期（一般不下于3年）、跨业务/跨地域的重大资源配置行动，其核心特征包括时间维度（分阶段长期规划）、空间维度（跨领域协同或地理扩张）、资源力度（重大资本投入或组织变革），例如业务结构调整、并购重组、产业链整合、海外建厂等，但这里不包含股权激励计划。示例：

- 公司推进昆药融合计划，业务聚焦三七产业链——跨业务整合
- 东重力储能项目顺利封顶、并网在即——跨周期产能部署
- 公司启动东南亚市场拓展计划，预计5年覆盖主要国家——跨地域扩张

技术创新：通过研发投入显著增加（如研发费用率 $\geq 5\%$ ）、技术突破（如专利授权、新产品上市）实现产品/工艺/商业模式的实质性升级。例如1类新药研发、AI工艺优化、区块链模式重构等。示例：

- 公司1类新药QBH-196进入I期临床试验——突破性产品创新
- 企业研发费用率提升至6%，新一代芯片性能提升40%——技术突破
- 公司区块链技术应用于供应链金融，提高交易效率——商业模式重构

渗透率提升：公司通过自主实施的渠道拓展/优化、产品创新/优化、服务升级等举措，直接推动市场份额扩大、认可度提升或客户消费频次/金额提升。核心特征是公司主动行为，且效果可量化，如市占率数据、客户数量增长、客单价提升等。示

例:

- 999 感冒灵认知度提升, 带动抗病毒口服液快速增长——知名度提升
- 虽然行业进入衰退期, 但公司通过渠道扩展提高了市占率——渠道拓展
- 品牌通过直播带货拓展线上渠道, 市占率提升 5%——渠道拓展
- 业优化产品功能, 客户复购率提升至 40%——产品优化

经营效率提升: 公司通过强化成本管控、周转率提升等运营过程的优化, 这里包含已体现在业绩的优化, 也包括已实施但效果尚未显示出来的优化措施, 如费用优化即使净利率未改善也算运营效率提升。示例:

- 库存周转率从 5 次提升至 7 次, 降低运营成本——已见效, 周转率提升
- 企业实施零基预算, 销售费用率从 15% 降至 12%——费用优化
- 公司引入自动化生产线, 人均产出提升 30%——生产效率提升

2.1.2 财务表现: 盈利质量、稳健性与业绩超预期分析

“财务表现”类重点关注公司经营成果, 已经发生而非分析师预期, 包括盈利改善、财务稳健性、业绩对比预期与同业是否存在超预期。

盈利改善: 关注“质量而非增速”。企业通过内生性经营优化(如产品结构升级、技术降本、产能释放), 实现核心盈利能力指标持续性提升, 如毛利率同比提升 ≥ 2 个百分点、营收/扣非净利润同比增长 $\geq 10\%$ 、ROE 同比提升 ≥ 1.5 个百分点等。示例:

- 毛利率提升 2.5 个百分点, 主因 AI 质检系统降本——内生性降本
- 扣非净利润同比增 12%, 得益于新技术降本增效——技术降本增效
- 营收增 10% 且毛利率稳定在 35%——稳健增长
- 营收增长 20% 但毛利率下降——低毛利扩张
- ROE 提升 1.8 个百分点至 16%, 产能利用率提高——产能释放

财务稳健性提升: 核心是验证是否降低系统性风险。企业通过资本结构优化(资产负债率同比下降 ≥ 3 个百分点)、债务控制(短期债务占比 $\leq 30\%$)、现金流改善(经营性现金流净额/净利润 $\geq 80\%$)等方式显著增强抗风险能力, 但需排除融资输血或资产出售等非经营手段, 且需多维度验证(如流动性、负债率、现金流同步优化)。示例:

- 经营性现金流净额同比增 51.16%, 负债率下降至 50%——多维优化
- 现金流改善但短期借款增加致负债率上升——单一改善但其他恶化
- 资产负债率降 3.2 个百分点, 经营现金流/净利润达 95%——财务健康增强
- 短期债务占比 25%, 现金短债比升至 1.5 倍——偿债能力提升

业绩对比预期和同业: 业绩对比预期与同业是超预期、低于预期还是符合预期。公司经营结果(如利润、营收、ROE、毛利率、订单、资产质量等)好于市场预期或分析师预测或行业平均水平。一般而言, 若指标表现优于预期或显著高于行业水平的关键词为“超预期、超预测、高于预期、超同业、超行业、创纪录、超指引”等; 若指标表现低于预期或显著差于行业水平, 关键词为“不及预期、低于预期、弱于同业、不及指引、逊于行业”等。示例:

- Q2 营收超预期增长, 净利润率达历史高点——超预期表现
- 毛利率达 40%, 显著高于行业均值 35%——超行业水平

- 净利润同比增长 25%，超分析师预期 15%——超预期增长
- ROE 18%，远高于同业平均 12%——行业领先
- 净利润同比增长 30%，基本符合预期——符合超预期

2.1.3 其他因素：股东行为、政策影响与市场驱动分析

“其他因素”类主要是公司外部因素，包括股东行为、政策驱动、市场驱动。

股东行为：包括股份回购（需区分注销或员工激励用途）、大股东及管理层增减持、股息政策调整（提高/降低分红、暂停派息）等。

- 药明康德拟斥资 10 亿元回购股份并注销——股份回购
- 控股股东通过二级市场增持公司股份至 28.5%，显示对公司前景的信心——大股东增持
- 公司决定提高年度股息率至 6%，以回馈长期投资者——股息政策调整
- x 高管减持所持股份的 15%，引发市场关注——减持套现
- 回购股份用于员工股权激励计划，以绑定核心人才——员工激励回购

政策驱动：由政府或监管机构直接针对公司或所在行业发布的政策，且会对公司未来的核心经营指标（收入/利润/成本等）产生可量化的直接影响，包含两类情形：(1)直接财务影响：政策直接改变公司收入、成本或现金流（如补贴、税收优惠、专项采购等）；(2)间接需求驱动：政策放松/收紧通过改变行业供需环境，导致公司未来需求变化（如放宽行业准入、刺激终端消费等）。

- 公司获高新技术企业认证，所得税率降至 15%——税收优惠
- 公司获新能源汽车购置补贴，单台补贴 2 万元，刺激销量增长——补贴支持
- 政策放松行业准入，公司计划新设 5 家分支机构——行业准入放宽
- 因环保政策趋严，公司需追加环保设备投资 1 亿元，短期内影响利润——成本增加
- 政府出台家电以旧换新补贴政策，公司产品销量预计提升 20%——市场需求刺激

市场驱动：反映市场因素对公司的直接影响，包括(1)资金驱动：反映外部机构资金行为对公司的持股情况，包括国家队、公募基金、保险资金、北向资金/陆股通、QFII、产业资本等市场机构投资者对公司持股比例的情况及潜力。(2)公司所在的赛道热度、景气度情况。

- 社保基金连续两个季度增持公司股份，持股比例升至 5%——机构资金流入
- AI 概念持续火热，公司 AI 芯片业务吸引多家机构调研——赛道热度提升
- 公募基金持仓比例下降至 12%，但北向资金增持至 8%——机构持仓变化
- 受行业龙头股价暴跌影响，公司股价当日跌幅达 5%——市场情绪影响
- 元宇宙概念退潮，公司相关业务板块估值回调 15%——赛道调整

2.2 摘要信息处理与多维度量分析

在本篇报告中，我们依托**慧博投研的分析师文本数据库**，对沪深 A 股自 2019 年 9 月 30 日至 2025 年 4 月 30 日的 38 万余篇研报摘要数据进行了系统处理。通过对

海量研报摘要的深度挖掘与分析，我们旨在从文本中提取关键信息，并从以下四个维度进行量化分析，以揭示研报中的潜在价值信息：

(1) **类别识别**：识别研报摘要中所涉及的利好类别，力求精准地梳理出各类潜在的积极因素，涵盖前边提到的 3 大类 10 小类，如战略布局、技术创新、盈利改善等。

(2) **情感倾向**：为便于快速统计与高效信息整合类别识别的结果，我们引入情绪倾向判断机制，即用简洁明了的数字标签来表征各类别所对应的情绪色彩。其中，1 代表正面利好，意味着该因素大概率对相关标的产生积极助推作用；-1 则象征负面利空，暗示存在潜在风险或不利影响；0 表示中性，即倾向性不明显或者影响程度微乎其微。

(3) **出现顺序**：在信息处理进程中，我们还融入了顺序记录，捕捉各类别在摘要中的出现次序。这一设计基于一个假设：前面提及的利好因素，是否是因撰写者的刻意安排或信息重要性排序，可能更具代表性与优先级，或是更能契合当下市场关注焦点，所以在处理过程中对该信息予以记录。

(4) **字数占比**：计算每个类别在摘要全文中的字数占比，这一指标能反映各类别在摘要中的文本权重。字数占比高的类别，通常包含更丰富的内容和细节，我们是否可参考其占比来判断该类别在核心观点中的重要性，从而聚焦关键信息呢？

2.3 模型选择与场景规则化训练

为了满足不同场景下的需求，我们对多种模型进行了测试和评估，并探索了本地部署与云端调用的可行性。以下内容将详细介绍我们在模型选择、任务规则化以及调用实践方面的具体工作。

2.3.1 模型选择与部署调用

根据前边的任务要求，我们分别测试了三种不同的模型：DeepSeek-R1-32B、DeepSeek-R1-671B、DeepSeek-V3-671B。

图表 4 不同尺寸的 DeepSeek-R1 版本介绍、参数量、特点、使用场景和硬件配置

DeepSeek模型版本	参数量	特点	适用场景	硬件配置
DeepSeek-R1-1.5B	1.5B	轻量级模型，参数量少，模型规模小	适用于轻量级任务，如短文本生成、基础问答等	4核处理器、8G内存，无需显卡
DeepSeek-R1-7B	7B	平衡型模型，性能较好，硬件需求适中	适合中等复杂度任务，如文案撰写、表格处理、统计分析等	8核处理器、16G内存，Ryzen7或更高，RTX 3060 (12GB)或更高
DeepSeek-R1-8B	8B	性能略强于7B模型，适合更高精度需求	适合需要更高精度的轻量级任务，比如代码生成、逻辑推理等	8核处理器、16G内存，Ryzen7或更高，RTX 3060 (12GB)或4060
DeepSeek-R1-14B	14B	高性能模型，擅长复杂的任务，如数学推理、代码生成	可处理复杂任务，如长文本生成、数据分析等	i9-13900K或更高、32G内存，RTX 4090 (24GB)或A5000
DeepSeek-R1-32B	32B	专业级模型，性能强大，适合高精度任务	适合超大规模任务，如语言建模、大规模训练、金融预测等	Xeon 8核、128GB内存或更高，2-4张A100 (80GB)或更高
DeepSeek-R1-70B	70B	顶级模型，性能最强，适合大规模计算和高复杂任务	适合高精度专业领域任务，比如多模态任务预处理。这些任务对硬件要求非常高，需要高端的CPU和显卡，适合预算充足的企业或研究机构使用	Xeon 8核、128GB内存或更高，8张A100/H100 (80GB)或更高
DeepSeek-R1-671B	671B	超大规模模型，性能卓越，推理速度快，适合极高精度需求	适合国家级 / 超大规模 AI 研究，如气候建模、基因组分析等，以及通用人工智能探索	64核、512GB或更高，8张A100/H100
DeepSeek-V3	671B	通用对话模型，相较于DeepSeek-R1-671B，没有推理功能	注重多轮对话能力、复杂问题解决和跨领域知识覆盖，擅长处理开放域问答、逻辑推理等	Xeon 8核、128GB内存或更高，8张A100/H100 (80GB)或更高

资料来源：华安证券研究所整理

在本机（搭载 RTX 4090 32G 显卡）上，我们部署了 DeepSeek-R1-32B 模型，可直接通过程序本地调用。而 DeepSeek-R1-671B 和 DeepSeek-V3-671B 模型由于本地部署对硬件要求较高，我们选择通过 API 调用进行测试。

我们对多家 API 进行了反复对比，包括官方 API。从收费情况来看，各家 API 基本一致。综合考虑稳定性和响应速度后，我们最终选择了火山引擎提供的 API。

图表 5 测试模型的部署方式和调用费用

模型名称	使用方式	输入计费	输出计费
DeepSeek-R1-32B	本地部署		
DeepSeek-R1-671B	火山引擎API	¥ 4/百万tokens	¥ 16/百万tokens
DeepSeek-V3-671B	火山引擎API	¥ 2/百万tokens	¥ 8/百万tokens

资料来源：华安证券研究所整理

2.3.2 任务规则化：为大模型在投资领域“降噪”增效

在投资场景，精准与可靠的信息分析至关重要。然而，大语言模型常常因“幻想”问题（如事实性幻觉与忠实性幻觉）而输出不准确或不切实际的内容，这在严谨的投资决策场景中显然是不可接受的。为了确保模型输出能够尽量契合金融分析的需求，任务规则化显得尤为关键。通过明确的规则化设计，我们可以有效约束模型的输出，避免其“幻想”，从而显著提升模型的输出效率和可用性。

1. 核心任务规则化

目前大模型最大的问题是“幻想”（即事实性幻觉与忠实性幻觉），在金融领域这类严谨任务场景中，需要对任务提示进行明确规则化，通过结构化约束与量化标准

实现分析过程的确定性控制。该体系基于以下核心设计原则：

(1) **分类识别标准化(类别)**：基于预定义的 10 类研报类别，要求模型执行分类识别时严格遵循三要素原则：

- 仅输出文本中实际提及的类别
- 排除未显式出现的分类项
- 禁止自主扩展分类边界

参考提示：识别文本提到的内容类别(从预定义的 10 个类别中选择)，不必输出未提到的类别。

(2) **情绪评分确定性(情绪值)**：鉴于每个类别的情绪判断(利好、利空、中性)标准存在差异，为确保大批量任务中不同批次的情绪判断标准统一，避免出现评分标准漂移风险，需对每个类别的情绪进行基本说明与定义。

- 基础层：建立数值映射标准(1=利好, -1=利空, 0=中性)
- 规则层：在分类定义中融入情绪判断依据，尽可能消除模型主观判断空间

参考提示：根据“类别说明”模块判断每个类别的情绪倾向，1 表示正面利好，-1 表示负面利空，0 表示中性。

图表 6 研报摘要的利好点分类定义与情绪评分规则

类别号	类别	定义
业务发展 (公司优化举措, 不一定已经体现在披露业绩)		
1	战略布局	基于外部环境研判与内部资源评估, 企业主动进行或者制定的战略调整, 尤其是跨周期、跨业务/地域的重大资源配置行动, 例如业务结构调整、并购重组、产业链整合、海外建厂等, 不含股权激励计划。
2	技术创新	通过研发投入显著增加(研发费用率≥5%)、技术突破(专利授权、新产品上市)实现产品/工艺/商业模式的实质性升级。例如 1 类新药研发、AI 工艺优化、区块链模式重构等。
3	渗透率	公司通过自主实施的渠道拓展/优化、产品创新/优化、服务升级等举措, 直接推动市场份额扩大或客户消费频次/金额提升。核心特征是公司主动行为, 且效果可量化(如市占率数据、客户数量增长、客单价提升等)。
4	经营效率	公司通过强化成本管控、周转率提升等运营过程的优化, 优化措施已实施即使效果未显也算利好, 如费用优化即使净利率未改善也算运营效率提升。
财务表现 (公司经营成果, 已经发生而非分析师预期, 重点关注经常项目而非一次性项目)		
5	盈利改善	关注“质量而非增速”。企业通过内生性经营优化, 实现核心盈利能力指标持续性提升, 包括毛利率、营收/扣非净利润、ROE 的同比/环比提升等。
6	财务稳健性	核心是验证是否降低系统性风险。企业通过资本结构优化、债务控制、现金流改善等方式显著增强抗风险能力, 但需排除融资输血或资产出售等非经营手段, 且需多维度验证(如流动性、负债率、现金流同步优化)。多维度达标(正面利好 1); 单一改善但其他恶化(中性 0); 未达标(负面利空 -1)
7	业绩对比预期和同业	业绩对比预期与同业是否存在超预期。当文本明确提及公司经营结果(如利润、营收、ROE、毛利率、订单、资产质量等)与市场预期、分析师预测或行业平均水平的对比时, 按以下规则分类: (1) 正面利好 (+1): 指标表现优于预期或显著高于行业水平。关键词: 超预期、超预测、高于预期、超同业、超行业、创纪录、超指引等。(2) 负面利空 (-1): 指标表现低于预期或显著差于行业水平。关键词: 不及预期、低于预期、弱于同业、不及指引、逊于行业。(3) 中性 (0): 明确表示符合预期预期或与行业水平相当。
其他因素		
8	股东行为	包括股份回购(需区分注销或员工激励用途)、大股东及管理层增减持、股息政策调整(提高/降低分红、暂停派息)等。情绪判断需结合行为动机与财务影响。
9	政策驱动	由政府或监管机构直接针对公司或所在行业发布的政策, 且会对公司未来的核心经营指标(收入/利润/成本)产生可量化的直接影响, 包含两类情形: (1) 直接财务影响: 政策直接改变公司收入、成本或现金流(如补贴、税收优惠、专项采购等); (2) 间接需求驱动: 政策放松/收紧通过改变行业供需环境, 导致公司未来需求变化(如放宽行业准入、刺激终端消费等)。
10	市场驱动	(1) 资金驱动: 反映外部机构资金行为对公司的持股情况, 包括国家队、公募基金、保险资金、北向资金/陆股通、QFII、产业资本等市场机构投资者对公司持股比例的情况及潜力。(2) 公司所在的赛道热度、景气度情况。

资料来源：华安证券研究所整理

(3) **出现顺序**：涉及的类别在文本中首次出现的位置。

参考提示：记录每个类别在文本中首次出现的顺序(1, 2, 3, ...)

(4) **篇幅计量科学化(字数占比)**：指类别字数/摘要总字数，鉴于可能存在同类别内容跨段落或某段落涉及多个类别等情况，需细化说明。

- 计算公式：类别字数/摘要总字数(0.00-1.00)

- 复合文本处理：执行字级拆分与权重分配
- 容错机制：允许未定义分类存在，因此最终涉及的分篇篇幅总和可能小于文本总量

参考提示：该类别字数/摘要总字数，跨类别文本按字数比例拆分，优先级：公司行为（类别 1-4）>外部因素（类别 8-10）>财务表现（类别 5-7），无法拆分时归入最相关类别（取值范围 0~1，保留两位小数）；由于摘要可能包含未列举的分类，因此涉及类别的篇幅占比总和可能小于 1。

2. 输入、输出格式

输入格式：在输入格式中，system 字段需包含系统提示以明确定义任务及处理要求，user 字段接收清洗后的研报摘要文本作为输入数据。

- system: 系统提示，定义任务要求及处理规则。
- user: 输入内容，这里是指清洗后的研报摘要文本。

输出格式：输出格式要求以 JSON 数组形式呈现分析结果，其中每个数组元素对应一个提及的研报要素类别，包含"类别"、"情绪"、"出现顺序"、"字数占比"四个字段。对于测试阶段的 R1 模型，需同时输出思考过程与处理逻辑，以便验证任务执行合理性；而在生产环境全量任务执行时，为提升处理效率，仅保留结果字段（content），不再输出推理内容（reasoning_content）。示例：

```

[{"类别": 4, "情绪": 1, "出现顺序": 1, "篇幅占比": 0.5},
 {"类别": 2, "情绪": -1, "出现顺序": 2, "篇幅占比": 0.2},
 {"类别": 8, "情绪": 0, "出现顺序": 3, "篇幅占比": 0.3}
]
    
```

3. 完整提示

综上，我们汇总各个要素形成完整的任务提示：

图表 7 系统提示

任务说明

忘记之前的提示，你现在是一个专业的研报内容分类及情绪（正面利好/负面利空/中性）识别专家，负责对分析师撰写的上市公司研报摘要所提及的分类进行分析。

业绩上重点关注已实现盈利，而非分析师盈利预测；忽略风险提示/风险/风险因素/风险分析部分的内容。

请严格遵循下述要求完成任务，并按照指定格式输出结果（输出格式示例）。

输入

- **system**：系统提示，定义任务要求及处理规则。
- **user**：研报摘要文本。

主要任务

请独立地对每条研报摘要文本进行以下分析：

1. ****类别****：识别文本提到的内容类别（从预定义的 10 个类别中选择），不必输出未提到的类别。
2. ****情绪****：判断每个类别的情绪倾向，1 表示正面利好，-1 表示负面利空，0 表示中性。
3. ****出现顺序****：记录每个类别在文本中首次出现的顺序（1, 2, 3, ...）。
4. ****字数占比****：该类别字数/摘要总字数，跨类别文本按字数比例拆分，优先级：公司行为（类别 1-4）>外部因素（类别 8-10）>财务表现（类别 5-7），无法拆分时归入最相关类别（取值范围 0~1，保留两位小数）；由于摘要可能包含未列举的分类，因此涉及类别的篇幅占比总和可能小于 1。

类别说明(类别相互独立, 请独立分析)

{class_str} (注: 图表6 中的内容)

输出要求

1. ****格式****:
 - JSON 数组, 数组元素为提到的类别信息。
2. ****内容****:
 - 仅包含提到的类别, 字段包括: 类别、情绪、出现顺序、字数占比。
 - 未涉及的/不确定的类别不需要做任何陈述, 也不必出现在结果中。
3. ****字段说明****:
 - 类别: 对应类别编号 (1-10)。
 - 情绪: 1 (正面利好)、-1 (负面利空)、0 (中性)。
 - 出现顺序: 类别在文本中首次出现的位置。
 - 字数占比: 该类别字数/摘要总字数, 保留两位小数 (跨类别文本按比例拆分)。

示例

输入格式示例:

- system: [{'role': 'system', 'content': 系统提示}]
- user: [{'role': 'user', 'content': 摘要}]

输出格式示例:

```

` ` json
[{"类别": 4, "情绪": 1, "出现顺序": 1, "篇幅占比": 0.5},
 {"类别": 2, "情绪": -1, "出现顺序": 2, "篇幅占比": 0.2},
 {"类别": 8, "情绪": 0, "出现顺序": 3, "篇幅占比": 0.3}
]
` `
    
```

资料来源: 华安证券研究所整理

2.3.3 模型调用实践: 本地与云端的高效对接

我们通过 Python 实现了本地和云端两种调用方式, 以完成不同场景下的测试。

(1) 本地部署调用 (通过 Ollama 部署)

对于本地部署的模型, 我们选择了 DeepSeek-R1-32B 版本, 利用 Ollama 框架进行高效调用。具体代码如下:

```

from ollama import chat

# 输入信息
messages = [{'role': 'system', 'content': prompt_system}]
messages.append({'role': 'user', 'content': f"{cleaned_text}"})

# 模型调用
Model = 'deepseek-r1:32b'
response: ChatResponse = chat(model=Model, messages=messages)
res_content = response.message.content
    
```

(2) 云端 API 调用 (通过 OpenAI 接口调用火山引擎 API)

对于硬件要求较高的模型, 如 DeepSeek-R1-250120, 我们选择通过 API 调用云端资源。在对比多家 API 服务后, 我们选择了火山引擎提供的 API, 其稳定性和响应速度均能满足我们的需求。具体代码如下:

```

# 初始化客户端
client = OpenAI(
    api_key=API_KEY,
    base_url="https://ark.cn-beijing.volces.com/api/" # 火山引擎API地址
)

# 输入信息
messages = [{'role': 'system', 'content': prompt_system}]
messages.append({'role': 'user', 'content': f"{cleaned_text}"})

# 模型调用
Model = "deepseek-r1-250120" # 以r1为例
response = client.chat.completions.create(
    messages=messages,
    model=Model,
    temperature=0.1,
    stream=False
)

# 获取结果
res_content = response.choices[0].message.content # 结果部分
res_reasoning_content = response.choices[0].message.reasoning_content # 推理部分
    
```

3 基于 DeepSeek 的分类任务实践与企业价值洞察

3.1 DeepSeek 在研报分类任务的实践

在本研究的研报利好分类任务中，我们的目标是通过系统提示规则，对文本内容进行归类、情绪判断、记录出现顺序、计算字数占比，并按照规定格式输出结果。这一任务的复杂性要求模型具备高准确率、稳定性以及对任务规则的严格遵循能力。为了评估不同模型在这一场景下的表现，我们对三种模型（DeepSeek-R1-32B、DeepSeek-R1-671B、DeepSeek-V3-671B）进行了详细测试，重点关注分类结果的合理性、输出格式的正确率以及结果的稳定性。

DeepSeek-R1-32B 模型的推理过程基本符合逻辑，但在输出结果格式的准确率方面表现欠佳，仅为约 50%。另外，多次测试同一输入时，其分类稳定性也未能达到理想水平。这表明该模型在处理复杂任务时，可能因参数量较小而难以完全适应此复杂场景的任务要求。

图表 8 DeepSeek-R1-32B 模型调用的输入输出示例：输出格式不符合规则，不是 JSON 格式

输入	推理过程	分类结果
<p>核心观点： 收入下降业绩承压，利润率及减值影响盈利能力。根据企业年报，23年万科实现营业收入4657.4亿元，同比下降7.6%，归母净利润121.6亿元，同比下降46.4%，受表内外利润率下降及资产减值影响，公司开发业务盈利能力承压。股东回报方面，受市场景气等因素影响，公司判断未来经营不确定性加大，为留足资金应对市场挑战，公司决定23年不派股息、不送红股。</p> <p>销售仍属第一梯队，拿地力度居主流房企中游。销售方面，23年公司实现销售3761亿元，同比下降10%，全国排名升至第2，维持第一梯队的规模水平。拿地方面，23年拿地总金额849亿元，对应金额口径拿地力度23%，在持续经营房企中力度处于中游。年末公司未售土储面积降至6851万方，同比下降19%，对销售覆盖倍数为2.78x，考虑资质厚度、结构及计划开工速度，未来公司销售去化一定程度上承压。</p> <p>有息负债规模稳定，主动规划调整负债结构。截至23年末公司账面上有息负债3201亿元，同比增长2%，其中银行贷款占比提升，公开债占比下降。公司在年报中提到，未来2年削减付息债务1000亿元以上，主动融入城市房地产融资协调机制，通过负债结构调整应对偿债压力。</p>	<p><think> 好，我现在要分析这篇研报摘要，并按照要求进行分类和情绪判断。首先，我得仔细阅读文本，理解每个段落的主要内容。 第一部分提到收入下降，业绩承压，利润率和减值影响盈利能力。xxxxx。 第二部分讨论销售情况，万科销售额3761亿元，同比下降10%，排名升至第2，维持第一梯队。xxxxxxx。 第三部分提到有息负债规模稳定，负债结构调整，未来计划削减债务，应对偿债压力。xxxxxxx。 现在逐一分析： 第一部分：xxxxxxx。 第二部分：xxxxxxx。 第三部分：xxxxxxx。 现在回到分类： xxxxxxxxxxxxx 因此，最终分类为： 1. 第8类股东行为，情绪中性。 2. 第10类市场驱动，情绪正面。 检查出现顺序：xxx。 字数占比方面，整篇摘要约450字左右： - 股东回报部分大约占总字数的三分之一。 - 销售情况部分也占三分之一。 - 剩下的是负债调整和其他内容，可能不涉及分类。 因此，假设每个提到的部分各占约30%的比例。但为了准确起见，需要更精确地计算： 第一部分（股东回报）：约150字 第二部分（销售情况）：约120字 第三部分（负债调整）：约180字 总字数约450字。 因此，第8类占比约为150/450≈0.33 第10类占比约为120/450≈0.27 负债调整部分可能不涉及分类，所以不需要计入。 但为了完整性，可以将剩余的部分分配到其他类别或忽略。因此，最终确定： - 第8类：33% - 第10类：27% 出现顺序为第8类在先，然后是第10类。 综上所述，任务的分类结果应该是： 1. 第8类股东行为，情绪中性。 2. 第10类市场驱动，情绪正面。 </think></p>	<p>#### 分类与分析</p> <p>1. **股东回报** - **类别**：第8类（股东行为） - **情绪**：中性 - 万科决定不分红，尽管通常被视为负面信号，但公司解释是为了应对经营不确定性并保留资金以支持业务发展，因此整体情绪趋向中性。</p> <p>2. **销售与市场地位** - **类别**：第10类（市场驱动） - **情绪**：正面 - 尽管销售额同比下降10%，但万科在行业中维持了第一梯队的地位，并且排名提升至第二位，显示出公司在市场中的竞争力和稳定性。</p> <p>3. **负债调整与财务策略** - 万科计划削减债务并优化负债结构，这是公司内部财务管理的决策，主要目的是为了应对偿债压力和改善财务状况。虽然不属于明确分类，但反映了公司的稳健经营策略。</p> <p>#### 结论</p> <p>- **第8类（股东行为）**：情绪中性，涉及不分红以保留资金应对市场挑战。 - **第10类（市场驱动）**：情绪正面，万科在销售和市场地位上表现稳定，维持了第一梯队的地位。 这些分类反映了万科当前的经营状况、财务策略以及市场定位。</p>

资料来源：慧博投研，华安证券研究所整理

DeepSeek-R1-671B 模型的推理过程严谨且符合逻辑，输出格式的准确率接近 100%，显示出其对任务规则的良好理解。然而，尽管其分类结果在明确类别上表现准确，但在易混淆的类别中，多次测试仍可能出现不同结果。这可能是由于模型的弱规范约束和自主决策的融入导致路径灵活性增加，从而引入了结果的多样性。

图表 9 DeepSeek-R1-671B 模型调用的输入输出示例：输出格式符合规则

输入	推理过程	分类结果
<p>核心观点： 收入下降业绩承压，利润率及减值影响盈利能力。根据企业年报，23年万科实现营业收入4657.4亿元，同比下降7.6%，归母净利润121.6亿元，同比下降46.4%，受表内外利润率下降及资产减值影响，公司开发业务盈利能力承压。股东回报方面，受市场景气等因素影响，公司判断未来经营不确定性加大，为留足资金应对市场挑战，公司决定23年不派股息、不送红股。</p> <p>销售仍属第一梯队，拿地力度居主流房企中游。销售方面，23年公司实现销售3761亿元，同比下降10%，全国排名升至第2，维持第一梯队的规模水平。拿地方面，23年拿地总金额849亿元，对应金额口径拿地力度23%，在持续经营房企中力度处于中游。年末公司未售土储面积降至6851万方，同比下降19%，对销售覆盖倍数为2.78x，考虑资质厚度、结构及计划开工速度，未来公司销售去化一定程度上承压。</p> <p>有息负债规模稳定，主动规划调整负债结构。截至23年末公司账面上有息负债3201亿元，同比增长2%，其中银行贷款占比提升，公开债占比下降。公司在年报中提到，未来2年削减付息债务1000亿元以上，主动融入城市房地产融资协调机制，通过负债结构调整应对偿债压力。</p>	<p>好的，我现在需要处理用户提供的万科年报摘要，并按照任务要求进行分类和情绪分析。首先，我要仔细阅读摘要内容，识别出涉及的类别，并确定每个类别的情绪、出现顺序和字数占比。</p> <p>首先看摘要的第一部分：xxxxx。 第二部分：“股东回报方面，受市场景气等因素影响，公司判断未来经营不确定性加大，为留足资金应对市场挑战，公司决定23年不派股息、不送红股。”这里提到不派股息，属于股东行为（类别8）。情绪是负面利空-1，因为停止派息通常视为负面。出现顺序是第二段，顺序可能是2。 第三部分：xxxxx。第四部分：xxxxx。第五部分：xxxxx。 现在需要确定每个提到的类别： 1. 盈利改善（类别5）：收入和利润下降，属于负面-1。出现顺序1。 2. 股东行为（类别8）：不派息，负面-1。出现顺序2。 3. 业绩对比同业（类别7）：销售排名升至第2，属于优于同业，情绪+1。出现顺序3。 4. 财务稳健性提升（类别6）：负债结构调整，可能属于中性0，因为负债规模稳定，但结构优化。出现顺序5。 5. 战略布局（类别1）或经营效率提升（类别4）：拿地力度和土储调整可能属于战略布局中的主动调整，但原文提到主动规划调整负债结构，可能属于战略布局。但负债结构调整属于财务稳健性，而拿地属于战略布局？需要看类别定义。战略布局是重大资源配置，如拿地属于战略布局吗？可能属于，但拿地力度中游，可能属于中性。或者可能属于经营效率提升中的周转率，但不确定。这部分可能需要进一步分析。 现在需要确定每个部分的字数占比： 摘要总字数需要计算，但用户提供的摘要可能已经分段。假设总字数约为400字左右，但需要精确计算。不过可能无法准确计算，但根据比例拆分。 例如，第一部分讨论盈利下降，属于类别5，字数约100字。股东行为部分约50字。销售部分约80字，属于类别7。拿地部分约70字，可能属于类别1或4。负债部分约100字，属于类别6。 但需要根据优先级拆分：公司行为（1-4）>外部因素（8-10）>财务表现（5-7）。如果有跨类别文本，按优先级拆分。 xxxxx 可能最终涉及的类别有： 类别5：盈利改善，情绪-1，顺序1，字数占比约25%。 类别8：股东行为，情绪-1，顺序2，字数占比约12.5%。 类别7：业绩对比同业，情绪+1，顺序3，字数占比约20%。 类别1：战略布局（拿地力度） 情绪0 顺序4 字数占比约17.5%。</p>	<pre> { "json": [{ "类别": 5, "情绪": -1, "出现顺序": 1, "字数占比": 0.25, "类别": 8, "情绪": -1, "出现顺序": 2, "字数占比": 0.13, "类别": 7, "情绪": 1, "出现顺序": 3, "字数占比": 0.2, "类别": 1, "情绪": 0, "出现顺序": 4, "字数占比": 0.18, "类别": 6, "情绪": 0, "出现顺序": 5, "字数占比": 0.24]] } </pre>

资料来源：慧博投研，华安证券研究所整理

DeepSeek-V3-671B 模型虽然没有实时推理过程，但其操作路径明确且标准化，生成目标的确定性高。测试结果显示，其分类结果与 R1 版本接近，且输出格式符合率高达 100%。在连续测试 100 条任务时，V3 版本的输出结果稳定性甚至略优于 R1

版本。此外，V3 版本支持批量请求，每次可输入多达 15 条摘要，显著提高了任务执行效率。

图 10 DeepSeek-V3-671B 模型调用的输入输出示例：输出结果与 R1-671B 相似

输入	推理过程	分类结果
<p>核心观点：</p> <p>收入下降业绩承压，利润率及减值影响盈利能力。根据公司年报，23年万科实现营业收入4657.4亿元，同比下降7.6%，归母净利润121.6亿元，同比下降46.4%，受表内外利润率下降及资产减值影响，公司开发业务盈利能力承压。股东回报方面，受市场景气等因素影响，公司判断未来经营不确定性加大，为留足资金应对市场挑战，公司决定23年不派发股息、不送红股。</p> <p>销售仍属第一梯队，拿地力度居主流房企中游。销售方面，23年公司实现销售3761亿元，同比下降10%，全国排名升至第2，维持第一梯队的规模水平。拿地方面，23年拿地总金额849亿元，对应金额口径拿地力度23%，在持续经营房企中力度处于中游。年末公司未售土储面积降至6851万方，同比下降19%，对销售覆盖倍数为2.78x，考虑货值厚度、结构及计划开工速度，未来公司销售去化一定程度上承压。</p> <p>有息负债规模稳定，主动规划调整负债结构。截至23年末公司账面有息负债3201亿元，同比增长2%，其中银行贷款占比提升，公开债占比下降。公司在年报中提到，未来2年削减付息债务1000亿元以上，主动融入城市房地产融资协调机制，通过负债结构调整应对偿债压力。</p>		<pre>```json [{"类别": 5, "情绪": -1, "出现顺序": 1, "字数占比": 0.25}, {"类别": 8, "情绪": -1, "出现顺序": 2, "字数占比": 0.15}, {"类别": 7, "情绪": 0, "出现顺序": 4, "字数占比": 0.20}], {"类别": 6, "情绪": 0, "出现顺序": 3, "字数占比": 0.20} ```</pre>

资料来源：慧博投研，华安证券研究所整理

综合来看，DeepSeek-R1 和 DeepSeek-V3 的 671B 满血版本均能基本完成给定任务，但 R1-32B 版本在复杂任务场景下表现欠佳。R1 版本由于存在推理过程，反应速度略慢于 V3 版本，且不支持批量任务请求，只能逐条输入“系统提示+1 条研报摘要”进行调用，效率较低。相比之下，DeepSeek-V3-671B 不仅支持批量请求，还具备更高的输出格式符合率和结果稳定性。

图 11 DeepSeek-V3-671B 模型基本符合本研究任务

模型名称	分类结果	输出格式	稳定性	响应速度	批量任务
DeepSeek-R1-32B	符合逻辑	50%符合	一般	单线程，5条/分钟	不支持
DeepSeek-R1-671B	符合逻辑	100%符合	较好	单线程5条/分钟	不支持
DeepSeek-V3-671B	符合逻辑	100%符合	稳定	单线程10条/分钟	支持

资料来源：华安证券研究所整理

因此，我们最终选择使用 DeepSeek-V3-671B 来完成本研究任务。为了进一步降低不同批次输入导致结果不稳定的风险，在正式执行任务时，我们对摘要的顺序进行了随机打乱。批量请求的输入格式如下：

批量请求的输入格式为：

role: system, content: 系统提示

role: user, content: 摘要 1

.....

role: user, content: 摘要 15

3.2 文本分类准确性的复查与调整

在完成所有研报摘要的处理工作后，我们对大模型分类结果进行了初步检查。过程中主要发现两类问题，并据此进行了针对性处理与优化。

(1) 模型无法输出预期结构化结果

部分摘要未能成功返回预期的 JSON 分类结构。针对这类情况，我们首先尝试

重新运行程序，判断是否属于偶发性错误。若多次运行仍无法解析，我们进一步排查原始摘要文本，发现多数问题源于文本中存在部分词汇，导致模型拒绝输出结果。对此，我们建立了词汇识别与替换机制，对相关文本进行适度修改后再次运行模型，从而恢复分类能力。

(2) 分类逻辑偏离传统“超预期”理解

在模型输出中，我们注意到**模型对分类 7 (业绩对比预期与同业) 的认定范围比我们最初设定的范围更为宽泛**。具体来看，该分类在全部研报摘要中的占比高达 21.9%。虽然模型对分类 7 的解释在某些情况下具有一定合理性，但为了使我们的分类方法与传统的“超预期”识别方法保持一致，我们决定对分类 7 进行进一步的优化处理。

图表 12 分类 7 (业绩对比预期与同业) 的分词逻辑

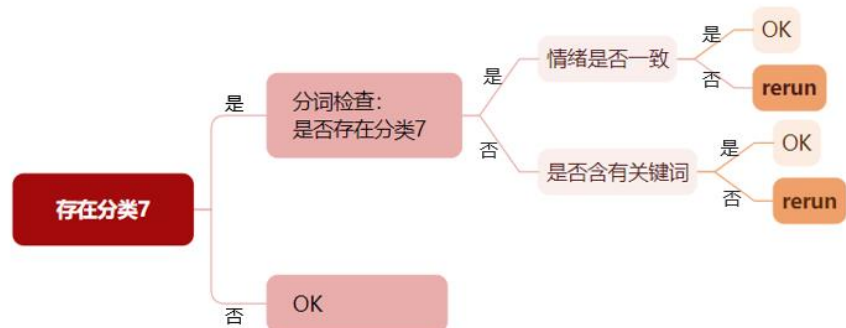
分词逻辑	
基础必选词组	预期, 预估, 预计, 承诺, 业绩预告, 预告上限, 预告, 一致预期, 市场预期, 市场一致预期, 指引, 行业水平, 行业, 同业
超预期必选词组	超, 超出, 超过, 高于, 好于, 优于, 大超
低于预期必选词组	低于, 不及, 小于, 不如, 逊于, 弱于
符合预期必选词组	符合, 介于, 处于
可选词组1	我们, 我们的, 我们此前, 我们此前的, 我们之前的, 我们在前瞻中的
可选词组2	略微, 略, 小幅, 大幅, 全面, 持续, 基本
屏蔽词组	料, 或, 将, 望, 有望, 可能, 待

资料来源：华安证券研究所整理

具体的复查逻辑如下：

- (1) 初步筛选：**检查模型初步输出结果是否包含分类 7。如果包含，则对该部分进行分词检查，以确保其准确性。
- (2) 情绪一致性检查：**如果分词后仍存在分类 7，进一步检查其情绪结果是否一致。如果情绪结果不一致，则重新运行模型进行分析。
- (3) 关键词检查：**如果分词后不存在分类 7，则检查文本中是否包含与分类 7 相关的关键词，如“行业平均”“同业”“行业水平”“预期”“承诺”“业绩预告”“预告上限”“预告”“一致预期”“市场预期”“市场一致预期”“创纪录”“超指引”等。如果文本中不包含这些关键词，则重新运行模型。

图表 13 分类 7 (业绩对比预期与同业) 的复查逻辑



资料来源：华安证券研究所整理

在优化分类 7 的过程中，我们发现了一个特殊情况：**部分研报摘要文本中并未涉及该分类，但模型返回的结果中却出现了分类 7 且情绪值为 0**。模型假设未明确

提及“超预期”或“不及预期”时，即为“符合预期”。然而，这种假设与其他分类逻辑及提示不符，其他分类基本遵循“未涉及则不输出”的原则。为了确保所有分类规则的一致性，我们决定剔除存在分类 7 且情绪为 0 的记录，并重新调整每篇报告中各分类的顺序。

3.3 细分类别助力从细节洞察企业真实价值

在本次金融文本分析任务中，分类识别作为核心环节，不仅有效避免了对研报摘要进行整体评分时出现的“笼统性”问题，还能更细致地揭示公司在不同经营维度上的具体表现。该分类方法的优势在于各类别相互独立，使我们能够从更全面且贴近主观研究的角度开展分析。以下是三种典型应用场景：

(1) 沙里淘金：洞察逆势价值，捕捉潜在反转

在整体业绩承压的背景下，通过分类识别挖掘局部亮点，预判企业的反转潜力。例如，即使公司盈利状况不佳，但通过运营优化和产品创新突破，公司仍可能具备实现反弹的潜力。如图表 14 中的示例，尽管公司盈利状况不佳，但通过运营优化和产品优势，公司具有实现反弹的潜力。这种分析方法和主观分析逻辑相契合，如周期底部行业筛选、困境反转策略标的的挖掘等逆向投资策略下的价值发现。

图表 14 “沙里淘金”举例

摘要原文	分析结果
<p>2024年10月29日公司披露三季报，前三季度实现营业收入13.64亿元，同比+2.70%；归母净利润0.82亿元，同比-13.49%；扣非净利润0.81亿元，同比-1.25%。24Q3实现营业收入4.72亿元，同比-1.17%；归母净利润0.22亿元，同比-42.08%；扣非归母净利润0.23亿元，同比-23.54%。</p> <p>行业竞争激烈，公司短时业绩承压。分渠道来看，xxxx。小B方面公司依托品牌与渠道优势稳固基本盘，同时适当给与一定促销政策保证增长，在行业竞争激烈的背景下，龙头企业占据一定优势。从产品端看，公司目前拥有油条、蛋挞等大单品，同时新品梯队搭建完善，虽然短时公司在大客户端推新进度有所放缓，但是公司研发能力行业领先，未来仍有通过新品实现快速增长的能力。</p>	<p>类别5（盈利改善）： 判定依据：归母净利润同比-13.49%，Q3归母净利润-42.08%，扣非净利率同比-1.41pct，核心盈利指标未达标。 情绪判断：净利润持续下滑且费用侵蚀利润（负面利空-1）。 字数计算：涵盖营收/利润增速对比、费用率攀升及利润率收缩等描述（占比40%）。</p> <p>类别4（运营效率提升）： 判定依据：小B渠道通过促销政策稳定基本盘，费用投放力度加大优化市场渗透。 情绪判断：渠道策略调整体现运营优化（正面利好1）。 字数计算：小B端渠道管理及促销策略段落（占比30%）。</p> <p>类别3（渗透率提升）： 判定依据：油条/蛋挞大单品+新品梯队搭建。 情绪判断：产品储备完善，可促进市场份额增长（正面利好1）。 字数计算：产品结构及研发能力描述（占比20%）。</p>

资料来源：慧博投研，华安证券研究所整理

(2) 去伪存真：剥离非核心因素，聚焦内生经营质量

通过剔除并表效应、汇率变动及政策影响等非持续性干扰，聚焦企业内生经营质量，还原企业的真实盈利水平。例如，某公司因并购及汇率波动导致净利润亏损，但剔除异常因素后，核心业务表现稳健且增长显著，符合“盈利改善”的判断逻辑。

图表 15 “去伪存真” 举例

摘要原文	分析结果
<p>2024年8月29日，公司披露2024年度中期业绩：公司2024年上半年实现营收5.3亿元，同比增长58.6%；扣非后归母净利润为-0.2亿元，较上年同期转亏。2024年第二季度，公司营业收入为2.6亿元，同比增长56.8%；扣非后归母净利润为-0.1亿元，较上年同期转亏。</p> <p>经营分析 报告期内公司营收大幅增长主要系2023年11月公司收购智明软件，间接持股100%；报告期内公司并表智明软件2.2亿元收入以及0.1亿元净利润。此外，公司营收和净利润还受到报告期内日元贬值的影响，xxxx，剔除智明软件并购和汇率影响后的公司扣非净利润为1.0亿元，同比增长160.0%。</p>	<p>类别5（盈利改善）：</p> <ul style="list-style-type: none"> • 圈定依据：剔除并购及汇率影响后，扣非净利润同比增长160%，符合“内生性经营优化实现核心指标达标”。 • 情绪判断：内生性利润增速远超10%且可持续（正面利好1）。 • 字数计算：剔除异常因素后的经营质量分析（占比28%）。

资料来源：慧博投研，华安证券研究所整理

(3) 拨云见日：剖析盈利本质，明确核心竞争力

区分经常性收益与非经常性损益（如政府补贴、资产处置），防止盈利质量被虚高数据掩盖，确保企业间的可比性。如图表 16 展示，某公司预告的归母净利润看似优异，但扣除非经常性损益后实际盈利表现不尽如人意。

图表 16 “拨云见日” 举例

摘要原文	分析结果
<p>公司预告上市公司整体全年归母净利润19-21亿元，其中领益科技自身23-25亿元，同比增长15.3%-25.3%。</p> <p>公司预告实际报表业绩19-21亿元，其中含有9.5亿元的业绩补偿股票公允价值变动，子公司亏损以及计提帝晶商誉减值对公司报表业绩略有拖累。</p>	<p>类别5（盈利改善）：</p> <p>判断依据：归母净利润增长但9.5亿补偿收益占比超45%，非经常性损益显著。□</p> <p>情绪：0（中性），因盈利改善依赖非持续收益且质量存疑。</p> <p>字数占比：盈利数据及补偿收益描述占25%。</p>

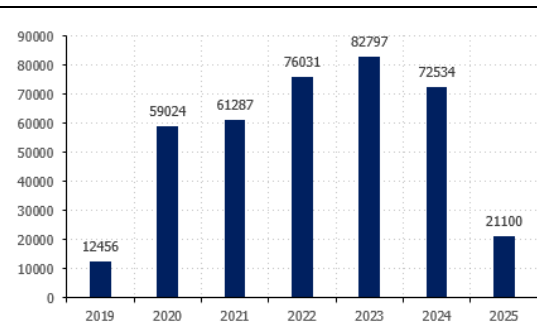
资料来源：慧博投研，华安证券研究所整理

4 结构化情绪视角下的研报摘要价值重构

4.1 研报分类与情绪分析全景

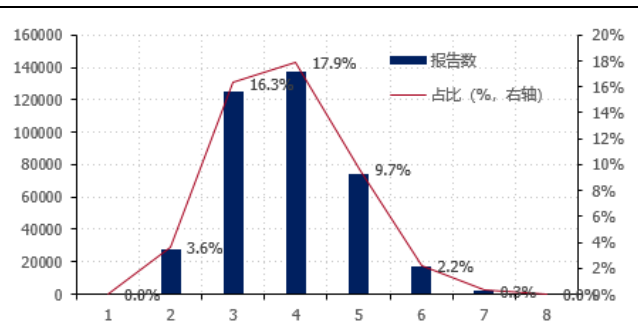
在统计区间 20190930 - 20250430 内，共有 38.5 万余篇报告获得了有效分类结果。从每篇报告所涉及的数量看，大部分报告涉及的类别数量为 3-5 类。

图表 17 分年度有效研报数量



资料来源：慧博投研，华安证券研究所

图表 18 研报涉及类别数统计

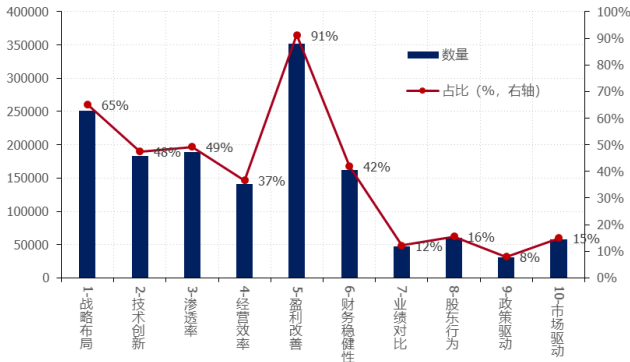


资料来源：慧博投研，华安证券研究所

总体来看，超过 90% 的报告涉及分类 5 - 盈利改善，占比（存在该分类的报告数/总报告数）最高；其次是分类 1 - 战略布局和分类 3 - 渗透率，分别约有 65% 和 49% 的报告涉及；而涉及比例最低的分类是分类 9 - 政策驱动、分类 7 - 对比

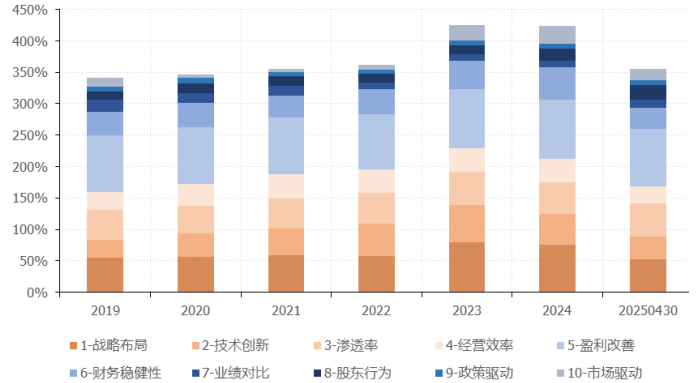
预期和同业和分类 10 - 市场驱动，占比分别为 8%、12% 和 15%。从年度分布来看，各年度各类别报告数量占当年报告总数的比例在时间序列上整体较为均衡。相对而言，分类 1 - 战略布局和分类 10 - 市场驱动（主要与行业板块赛道热度相关）被分析师提及的频率呈现逐年上升的趋势。另外，自去年以来，随着分红新规的实施，与分类 8 - 股东行为相关的公司行为日益增多，导致该分类的报告数量占比逐渐上升，近三年的占比分别为 13.6%、19.1% 和 23.2%。

图表 19 区间内，涉及各类别的报告数量与占比



资料来源：慧博投研，华安证券研究所

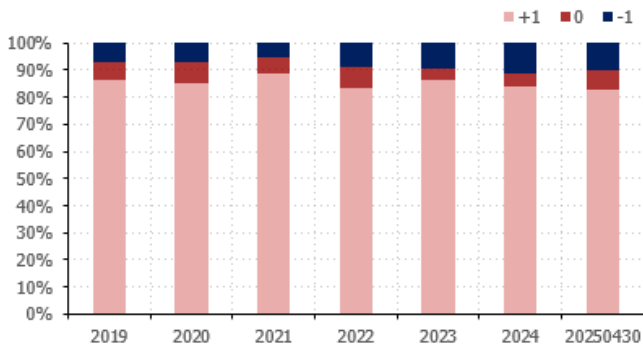
图表 20 分年度，各分类数量占报告数的情况



资料来源：慧博投研，华安证券研究所

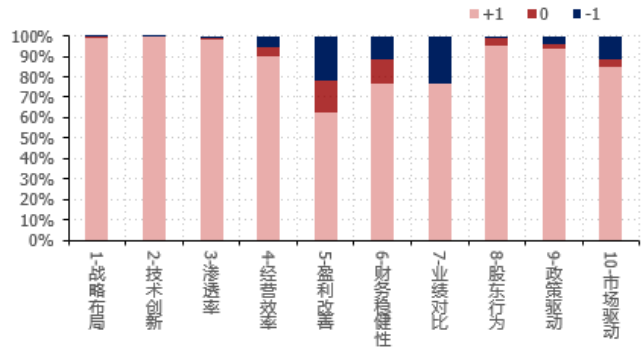
和预期一致，分析师在报告中倾向聚焦正向因素，对负向因素或不提，或持中性态度。从数据统计来看，所有报告的分类中，正向情绪 (+1) 占比（正向情绪数量/总情绪数量）高达 85.3%，中性(0)和负向(-1)情绪占比分别为 6.1% 和 8.6%。年度间的不平衡度略有差异，其中 2021 年和 2023 年表现得更为显著（图表 21）。从各分类的情绪分布来看（图表 22），所有分类均呈现出显著的正向情绪倾向。其中，分类 5 - 盈利改善因基于客观财务数据，情绪分布相对均衡一些，+1、0、-1 的情绪占比分别为 62.5%、15.3% 和 22.2%。

图表 21 各年度的情绪分布 (%)



资料来源：慧博投研，华安证券研究所

图表 22 各分类的情绪分布 (%)



资料来源：慧博投研，华安证券研究所

4.2 利好密度驱动下的收益预测与策略应用

在建立了情绪分类体系并完成情绪标签标准化后，我们尝试进一步探讨各类“利好情绪”在实际市场中的表现机制与定量效能。为此，我们以“未来超额收益”为核心

指标，分别从报告维度、个股维度以及选股策略构建维度，系统性分析“利好密度”与市场收益之间的关系，评估不同类别情绪对未来收益的指示性。

4.2.1 短期利好驱动强，超预期与盈利改善信息更具价值

我们首先从报告维度出发，分别测试每个类别的正负情绪与未来收益的关系。图表 23 展示了每个类别中，不同情绪的报告对应的未来累计收益 CAR 的相关性。这里，为了剔除不同时间的市场环境的影响，我们用相对全 A 的超额收益，观察时间窗口为 5、20、60 交易日。可以发现：

- (1) 在几乎所有类别下，情绪值为 +1 的报告的 CAR 值及 CAR 为正的的比例，明显高于情绪值为 -1 的报告，这表明市场对利好消息的交易通常比对利空消息更积极。
- (2) 当观察期较短（例如 5 日、20 日）时，利好的超额收益相比利空更为显著，这说明事件驱动型利好所带来的超额收益会随着事件的发酵和时间推移而逐步减弱。
- (3) 无论处于哪个时间窗口，分类 7 - 对比预期和同业（情绪值为 +1）的报告的 CAR 值和 CAR 为正的的概率均最高，显示出市场对业绩超预期这一明确利好信号的高度敏感性。
- (4) 鉴于其他分类的正样本占比远高于负样本，负样本的统计收益均值易受离群值影响。然而，分类 5 - 盈利改善的样本分布相对平衡。结果依然是情绪值为+1 的报告对应的 CAR 均显著高于情绪值为-1 的报告。表明市场对盈利改善类别下不同情绪的反应较为理性，且该类信息在不同时间维度上均具有一定的价值。

图表 23 各类别，不同情绪的报告对应的未来 CAR 情况

类名	情绪值	报告占比	超额收益均值 (%)			CAR为正的的比例		
			CAR5	CAR20	CAR60	CAR5	CAR20	CAR60
1-战略布局	-1	0.10%	-0.92	-0.31	-0.92	42.1%	43.4%	39.8%
	+1	64.48%	0.47	0.59	0.92	48.6%	46.4%	42.3%
2-技术创新	-1	0.04%	-0.26	-1.12	-1.23	42.7%	40.8%	41.4%
	+1	47.46%	0.51	0.66	1.61	48.8%	46.4%	43.0%
3-渗透率	-1	0.39%	-0.25	0.12	-0.23	42.9%	44.6%	39.7%
	+1	48.11%	0.57	0.73	1.19	49.1%	46.8%	41.9%
4-运营效率	-1	1.98%	-0.07	0.06	-0.70	45.5%	45.2%	40.1%
	+1	33.13%	0.55	0.64	0.82	49.3%	47.0%	42.4%
5-盈利改善	-1	20.26%	0.16	0.64	0.89	47.0%	47.5%	42.4%
	+1	57.04%	0.65	0.69	1.27	49.6%	46.7%	41.9%
6-财务稳健性	-1	4.85%	0.07	0.30	0.50	46.1%	46.0%	42.3%
	+1	32.22%	0.40	0.48	0.30	48.5%	46.7%	41.8%
7-业绩超预期	-1	2.86%	-0.34	0.23	0.12	43.2%	45.9%	40.2%
	+1	9.26%	1.33	1.42	2.82	54.4%	50.1%	45.4%
8-股东行为	-1	0.19%	0.16	0.08	-0.26	46.8%	45.1%	40.6%
	+1	15.00%	0.58	0.87	0.57	49.5%	48.3%	41.3%
9-政策驱动	-1	0.30%	0.22	1.04	-0.40	48.7%	49.9%	38.6%
	+1	7.59%	0.54	0.74	0.85	48.5%	46.6%	42.3%
10-市场驱动	-1	1.75%	0.12	0.36	0.16	45.9%	47.2%	42.1%
	+1	12.81%	0.37	0.43	0.56	48.2%	46.2%	42.0%

资料来源：wind 资讯，慧博投研，华安证券研究所

4.2.2 从报告到个股：分类利好密度可有效预测未来收益

前文在报告维度中验证了利好情绪与未来收益之间的关系：整体来看，利好类报告的累积超额收益（CAR）普遍高于利空类，符合预期。进一步地，我们将分析视角切换至个股维度，探讨在指定时间区间内，若某类利好消息占比越高，是否能带来更高的未来超额收益。为此，我们构建了如下指标：

某股票的某类别利好的占比

$$= \text{该股票在指定区间内, } \frac{\text{该分类情绪值之和}}{\text{该股票同期内所有情绪的总数量}}$$

该占比指标（CatSentiRatio）用于衡量特定时间段内，某只股票所有报告的某一类别的利好情绪在该股票在区间内的所有报告涉及的全部情绪中的相对占比，从而量化个股受到该类别利好“密度”的影响程度，后文称之为“利好密度”。具体计算步骤如下：

- (1) 确定报告时间 t：以每条分类情绪对应的报告发布时间为基准；
- (2) 计算该分类情绪和（CatSentiSum）：统计 t-N 个交易日内，个股的全部报告的该分类的情绪值总和；
- (3) 统计全部情绪数量（AllSentiNum）：同样在 t-N 个交易日内，统计个股的全部报告的所有情绪（无论正负）的数量；
- (4) 计算利好密度：即 CatSentiSum / AllSentiNum
- (5) 在此基础上，针对每一情绪类别，运行如下 OLS 回归模型，探索利好占比与未来收益之间的关系：

$$CAR5/CAR20/CAR60 = \beta_0 + \beta_1 * \text{利好密度} + \epsilon$$

如需控制行业影响（中信一级行业），模型扩展为：

$$CAR5/CAR20/CAR60 = \beta_0 + \beta_1 * \text{利好密度} + \beta_2 * \text{行业虚拟变量} + \epsilon$$

从图表 24-图表 25 的回归结果可见：大多分类中，利好密度的系数 β_1 均显著为正，表明利好密度与未来收益之间存在稳定的正向关系。尤其是“分类 7-对比预期和同业”最为显著，该类利好在情绪中占比越高，未来获得正向超额收益的概率越大。

图表 24 个股维度：各分类利好密度回归结果

利好类别	回归系数			系数显著性 (P-Value)		
	CAR5	CAR20	CAR60	CAR5	CAR20	CAR60
1-战略布局	0.872	-0.134	-1.221	0.0001	0.7533	0.1139
2-技术创新	1.617	3.319	8.372	0.0000	0.0000	0.0000
3-渗透率	3.350	5.131	8.313	0.0000	0.0000	0.0000
4-经营效率	2.225	2.584	2.537	0.0000	0.0000	0.0011
5-盈利改善	0.669	0.274	2.366	0.0000	0.0871	0.0000
6-财务稳健性	0.122	-0.988	-4.366	0.4948	0.0028	0.0000
7-业绩超预期	5.097	4.819	15.010	0.0000	0.0000	0.0000
8-股东行为	2.067	3.577	1.626	0.0000	0.0000	0.1023
9-政策驱动	2.706	4.694	6.869	0.0000	0.0000	0.0000
10-市场驱动	1.389	2.387	6.655	0.0000	0.0000	0.0000

资料来源：Wind 资讯，慧博投研，华安证券研究所

图表 25 个股维度：各分类利好密度回归结果（行业因素）

利好类别	回归系数			系数显著性 (P-Value)		
	CAR5	CAR20	CAR60	CAR5	CAR20	CAR60
1-战略布局	0.530	-0.947	-4.750	0.0268	0.0346	0.0000
2-技术创新	1.527	2.772	2.788	0.0000	0.0000	0.0005
3-渗透率	3.238	4.885	6.812	0.0000	0.0000	0.0000
4-经营效率	2.346	2.946	5.630	0.0000	0.0000	0.0000
5-盈利改善	0.637	0.072	1.633	0.0000	0.6577	0.0000
6-财务稳健性	0.525	-0.606	-1.965	0.0058	0.0846	0.0028
7-业绩超预期	5.133	4.714	13.998	0.0000	0.0000	0.0000
8-股东行为	2.416	4.095	5.297	0.0000	0.0000	0.0000
9-政策驱动	2.962	5.708	9.238	0.0000	0.0000	0.0000
10-市场驱动	1.357	2.207	4.772	0.0000	0.0001	0.0000

资料来源：Wind 资讯，慧博投研，华安证券研究所

4.2.3 构建基于情绪利好密度构建主动选股策略

在确认了“利好密度”在个股层面对未来收益具有稳定预测能力后，我们进一步从实操角度出发，尝试构建基于不同情绪类别的主动选股策略。具体而言，我们尝试基于“利好密度”，以“分类 7”（最显著）与“分类 5”（样本量最大、情绪分布最均衡）为例构建类别分域下的主动量化选股组合。

- (1) 在每个报告时间 t ，计算近 N 个交易日的该分类利好密度；
- (2) 在每个调仓周期，保留个股近 6 个月内最新的利好密度记录；
- (3) 从全市场中筛选利好密度最高的前 K 只个股，构成等权重组合；如存在相同占比，则优先选择在各自行业中利好密度排名靠前的个股。

策略回测参数设置如下：

- 样本池：全 A，剔除 ST、*ST、上市 3 个月内、市值位于后 20% 的股票，以及北交所上市的个股。
- 回测时间：2019.12.31-2025.05.30
- 手续费：双边千 3
- 基准：中证 800
- 组合数量：30 只、50 只；组合内个股权重：等权配置
- 调仓频率：按月调仓，或在每年研报密集发布月份（1、4、7、8、10 月）月底调仓
- 交易规则：调仓次日以成交均价买入，考虑涨跌停等实际交易约束。

1. “盈利改善密度”优选股票组合

我们以“盈利改善密度”（分类 5）为核心指标，构建基于不同参数的主动股票组合，并测试以下三类变量对组合表现的影响：

- 时间窗口长度（ $N=20$ vs $N=40$ ）
- 持股个数（Top30 vs Top50）
- 调仓频率（月频 vs 不定期）

图表 26-图表 28 为组合的回测结果，可以看到：**短期情绪密度信号更有效、集中持股优于分散持股、调仓频率影响有限，但月频回撤更低。**

- (1) **短期情绪密度信号更有效**： $N=20$ 优于 $N=40$ 。组合在使用近 20 日利好密度（ $N=20$ ）构建时，整体收益更高、回撤更小。这可能归因于“盈利改善”相关研报数量本身较多，若回取区间过长（ $N=40$ ），会引入更多滞后性信号，导致选股有效性下降。
- (2) **集中持股优于分散**：Top30 优于 Top50。在同一调仓频率下，选择利好密度排名前 30 的个股（Top30）比前 50 名（Top50）表现更佳。以月频调仓为例：“ $N=20$, Top30”组合：年化收益 15.0%，超额收益 15.6%，最大回撤 27.5%，“ $N=20$, Top50”组合：年化收益 10.8%，超额收益 11.4%，最大回撤 30.5%，表明密度信号在高排名区间的“边际有效性”更强。
- (3) **调仓频率影响有限，但月频回撤更低**。从不同调仓方式对收益端的影响来看，月频与不定期调仓在收益上差异不大，但月频组合的风险控制更优，回撤普遍更小。原因可能在于：盈利改善相关情绪信号密集，月频调仓更能及时响应新增研报，从而降低组合风险。

图表 26 盈利改善密度组合的分年度收益情况（月频调仓）

	中证800	N=40, Top30			N=40, Top50			N=20, Top30			N=20, Top50		
		收益率	收益率	超额收益	最大回撤	收益率	超额收益	最大回撤	收益率	超额收益	最大回撤	收益率	超额收益
2020	25.8%	29.6%	3.8%	-16.4%	30.5%	4.7%	-15.8%	29.2%	3.4%	-17.8%	25.4%	-0.4%	-18.9%
2021	-0.8%	43.0%	43.8%	-14.8%	30.3%	31.1%	-15.0%	36.1%	36.9%	-14.5%	28.9%	29.6%	-14.5%
2022	-21.3%	-12.0%	9.3%	-28.9%	-12.3%	9.0%	-26.0%	-10.0%	11.4%	-24.5%	-10.1%	11.2%	-25.0%
2023	-10.4%	6.5%	16.9%	-11.5%	2.6%	12.9%	-11.6%	10.4%	20.8%	-10.1%	5.2%	15.6%	-11.1%
2024	12.2%	8.5%	-3.7%	-20.7%	5.2%	-7.0%	-22.1%	8.5%	-3.7%	-19.7%	3.5%	-8.8%	-22.1%
20250530	-2.0%	7.9%	9.9%	-14.6%	3.4%	5.5%	-14.4%	9.6%	11.6%	-17.8%	8.0%	10.0%	-16.0%
整体	-0.6%	14.5%	15.1%	-30.9%	10.2%	10.8%	-33.3%	15.0%	15.6%	-27.5%	10.8%	11.4%	-30.5%

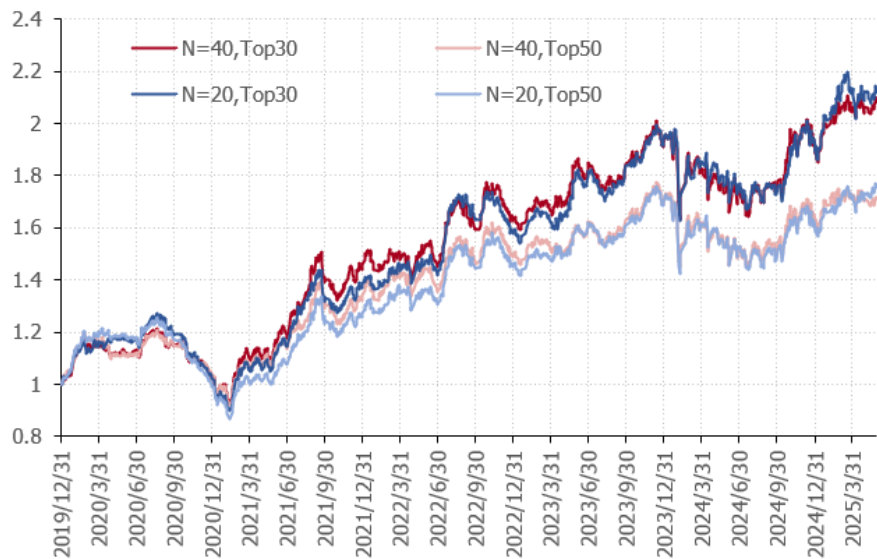
资料来源：wind 资讯，慧博投研，华安证券研究所

图表 27 盈利改善密度组合的分年度收益情况（不定期调仓）

	中证800	N=40, Top30			N=40, Top50			N=20, Top30			N=20, Top50		
		收益率	收益率	超额收益	最大回撤	收益率	超额收益	最大回撤	收益率	超额收益	最大回撤	收益率	超额收益
2020	25.8%	32.8%	7.0%	-16.7%	34.2%	8.5%	-16.7%	27.6%	1.8%	-14.7%	27.1%	1.3%	-17.0%
2021	-0.8%	31.6%	32.3%	-15.1%	18.3%	19.0%	-16.3%	27.7%	28.5%	-16.4%	15.1%	15.9%	-16.3%
2022	-21.3%	-16.6%	4.8%	-30.0%	-12.3%	9.0%	-26.4%	-10.1%	11.2%	-25.6%	-11.8%	9.5%	-27.3%
2023	-10.4%	20.0%	30.4%	-11.0%	6.8%	17.1%	-10.8%	20.4%	30.8%	-11.4%	11.9%	22.2%	-12.0%
2024	12.2%	5.3%	-6.9%	-23.8%	4.2%	-8.0%	-23.6%	10.9%	-1.4%	-22.0%	7.3%	-4.9%	-24.3%
20250530	-2.0%	7.5%	9.5%	-14.4%	1.1%	3.1%	-16.2%	7.6%	9.6%	-17.0%	6.8%	8.8%	-15.2%
整体	-0.6%	13.9%	14.5%	-33.5%	8.9%	9.5%	-34.1%	15.3%	15.9%	-32.1%	10.1%	10.7%	-33.1%

资料来源：wind 资讯，慧博投研，华安证券研究所

图表 28 盈利改善密度组合的超额净值曲线



资料来源：Wind 资讯，慧博投研，华安证券研究所整理

2. “业绩超预期密度”优选股票组合

采用同样的方法，我们构建以“业绩超预期密度”（分类 7）为核心指标的主动股票组合。结果显示：长期情绪密度信号更有效、调仓频率依然影响有限。

(1) 长期情绪密度信号更有效：使用近 40 日利好密度（N=40）构建的组合在收益和回撤控制方面均优于近 20 日（N=20）。可能因为“业绩超预期”类研报数量较少，短期样本不足易引入偶发性噪声，影响选股精度。这也导致 Top30 和 Top50 的组合表现差距不显著，样本稀疏时，过于集中可能

不稳定。

(2) 调仓频率依然影响有限：与“盈利改善密度”策略类似，月频与不定期调仓在收益表现上差异不大。

图表 29 业绩超预期密度组合的分年度收益情况（月频调仓）

	中证800	N=40, Top30			N=40, Top50			N=20, Top30			N=20, Top50		
	收益率	收益率	超额收益	最大回撤	收益率	超额收益	最大回撤	收益率	超额收益	最大回撤	收益率	超额收益	最大回撤
2020	25.8%	40.1%	14.3%	-12.9%	36.3%	10.5%	-14.2%	17.6%	-8.2%	-15.6%	24.1%	-1.7%	-14.4%
2021	-0.8%	18.1%	18.8%	-12.1%	21.8%	22.5%	-12.1%	24.8%	25.6%	-11.8%	19.4%	20.1%	-11.5%
2022	-21.3%	-2.7%	18.6%	-25.3%	-8.6%	12.7%	-26.0%	-11.4%	9.9%	-34.7%	-8.2%	13.2%	-29.2%
2023	-10.4%	9.5%	19.9%	-13.6%	1.6%	11.9%	-13.9%	0.3%	10.7%	-15.9%	6.4%	16.8%	-14.3%
2024	12.2%	-0.7%	-12.9%	-29.3%	-0.6%	-12.8%	-26.2%	-2.7%	-14.9%	-25.6%	-4.0%	-16.2%	-24.3%
20250530	-2.0%	2.5%	4.6%	-18.5%	5.2%	7.3%	-17.2%	7.0%	9.1%	-17.7%	5.4%	7.5%	-17.4%
整体	-0.6%	11.8%	12.4%	-33.3%	9.5%	10.2%	-34.0%	6.0%	6.6%	-35.7%	7.6%	8.2%	-29.5%

资料来源：wind 资讯，慧博投研，华安证券研究所

图表 30 业绩超预期密度组合的分年度收益情况（不定期调仓）

	中证800	N=40, Top30			N=40, Top50			N=20, Top30			N=20, Top50		
	收益率	收益率	超额收益	最大回撤	收益率	超额收益	最大回撤	收益率	超额收益	最大回撤	收益率	超额收益	最大回撤
2020	25.8%	46.8%	21.0%	-14.5%	45.2%	19.4%	-15.0%	28.7%	2.9%	-15.0%	36.1%	10.3%	-14.3%
2021	-0.8%	19.7%	20.4%	-13.5%	17.7%	18.5%	-12.9%	24.3%	25.1%	-11.4%	20.0%	20.7%	-12.2%
2022	-21.3%	-2.5%	18.9%	-21.2%	-9.5%	11.9%	-22.5%	-12.4%	8.9%	-32.8%	-8.9%	12.5%	-29.6%
2023	-10.4%	9.1%	19.5%	-14.7%	5.6%	16.0%	-13.4%	0.3%	10.6%	-14.8%	7.3%	17.7%	-13.8%
2024	12.2%	0.3%	-11.9%	-26.9%	4.6%	-7.6%	-23.9%	-4.9%	-17.1%	-26.9%	-4.4%	-16.6%	-26.9%
20250530	-2.0%	3.2%	5.3%	-19.4%	7.0%	9.0%	-16.6%	7.6%	9.7%	-17.8%	6.3%	8.4%	-16.9%
整体	-0.6%	14.2%	14.8%	-31.1%	12.9%	13.5%	-30.3%	7.2%	7.8%	-38.4%	9.7%	10.3%	-32.9%

资料来源：wind 资讯，慧博投研，华安证券研究所

此外，我们还可以结合投资研究经验，对情绪类别进行交叉组合。尽管每个单一类别的情绪倾向通常是单向的，但不同类别的组合却不再是简单的“积极”或“消极”叠加，而是呈现出多维度、复合化的情绪结构。这种结构化的情绪理解，有助于挖掘更细致的情绪信号，构建基于交叉情绪特征构建更具方向性的投资组合，提升组合的针对性与实用性。例如：

- 财务层面的乐观 vs 战略层面的保守
- 市场展望积极 vs 经营现状中性
- 短期信心高 vs 长期判断谨慎
- 业绩改善欠佳 vs 经营效率提高/战略布局/技术创新/渗透率提升

4.3 研报信息结构是否影响未来走势？

本小节我们进一步探讨研报中利好与利空信息的“强调方式”（包括篇幅比例与出现顺序）是否与未来超额收益存在关联。理论上，若某类利好信息被分析师优先呈现或详细阐述，往往表明其重要性较高，因而更可能推动股价上涨。相反，利空信息若被突出强调，则可能对后续收益产生负面影响。

4.3.1 顺序：类别信息靠前意味着更强定价效应？

我们首先考察利好/利空在摘要中的出现顺序是否影响后续收益。模型设定如下：

$$CAR5/CAR20/CAR60 = \beta_0 + \beta_1 * \text{标准化的出现顺序} + \beta_2 * \text{行业虚拟变量} + \epsilon$$

其中，“标准化的出现顺序”是指将原始顺序线性映射到(0,1]区间，即：

某报告某分类的 标准化的出现顺序

$$= \frac{\text{该报告的最大排序值} - \text{某分类的出现顺序原值} + 1}{\text{该报告的最大排序值}}$$

某一类情绪性信息在摘要中相对于其他内容出现的**靠前/靠后位置**，并进行标准化处理。理论假设是：**越靠前的信息，越可能反映分析师认为“最重要、最值得市场关注”的部分**，因此更可能对市场产生定价影响。核心结论如下：

- (1) **部分正向利好类别在摘要中出现顺序确实对后续收益具有一定解释力**。当某类利好信息在摘要中出现得更靠前，往往意味着该信息被分析师认为更为重要、更值得市场关注，从而更可能被投资者定价吸收。典型如 **3-渗透率**、**7-业绩对比预期和同业** 和 **8-股东行为**，在 CAR5、CAR20、CAR60 三个不同观察期中均展现出显著的正向收益预测能力。其中，“业绩超预期”是代表性因素，回归系数较大、显著性稳定，符合实际调研和写作经验：若公司业绩显著好于市场预期，分析师通常会将其作为核心亮点，安排在摘要开头突出强调。
- (2) **部分基础财务类因子（如经营效率、盈利改善、财务稳健性）反而呈现出负相关，顺序越靠前，未来收益越差**。这类现象一方面源于信息“惯性”：大多数研报为**业绩点评类**，出于写作逻辑，会优先对财务表现进行描述，并不完全代表分析师的关注重心；另一方面，也反映出分析师个人风格的差异。例如，一些报告习惯“先讲业绩，再讲逻辑”，另一些则更注重从战略、政策、行为等角度开篇切入。因此，这几类因子的出现顺序未必具有强解释力，甚至可能存在“越靠前反而是常规性描述”的结构性干扰。
- (3) **在负面情绪（-1）下，摘要顺序对收益基本没有解释力，且大多数类别的系数方向与预期不符**。理论上，利空因素若出现在摘要更靠前的位置，应该意味着其对公司负面影响更大，进而应对后续收益构成负向影响，即回归系数应为负，才具有预期内的预测能力。从结果看，没有任何分类表现出方向正确、系数显著的情况，甚至出现“越靠前的利空，收益越好”的反常情形（如“技术创新”，系数正向显著）。究其原因，一是分析师在表达负面信息时较为保守，常放在摘要末尾，导致排序信号天然失真；二是样本较小或研报写作风格差异，干扰排序变量的有效性。

图表 31 出现顺序与 CAR 的回归结果（情绪值为+1）

利好类别	回归系数			系数显著性 (P-Value)		
	CAR5	CAR20	CAR60	CAR5	CAR20	CAR60
1-战略布局	0.258	0.045	-0.652	0.0000	0.6703	0.0006
2-技术创新	0.223	0.623	-0.688	0.0050	0.0000	0.0086
3-渗透率	0.532	0.763	0.788	0.0000	0.0000	0.0021
4-经营效率	-0.048	-0.538	-0.916	0.6296	0.0031	0.0056
5-盈利改善	-0.723	-1.091	-0.542	0.0000	0.0000	0.0332
6-财务稳健性	-0.239	-0.117	0.746	0.0039	0.4306	0.0070
7-业绩对比	0.805	0.490	0.953	0.0000	0.0903	0.0790
8-股东行为	0.276	0.553	1.440	0.0146	0.0075	0.0001
9-政策驱动	0.266	0.527	1.197	0.1541	0.1142	0.0371
10-市场驱动	0.072	-0.234	0.328	0.6867	0.4773	0.5832

资料来源：Wind 资讯，慧博投研，华安证券研究所

图表 32 出现顺序与 CAR 的回归结果（情绪值为-1）

利好类别	回归系数			系数显著性 (P-Value)		
	CAR5	CAR20	CAR60	CAR5	CAR20	CAR60
1-战略布局	0.908	-2.507	-1.177	0.4724	0.4212	0.7855
2-技术创新	5.244	10.260	22.519	0.0377	0.0368	0.0092
3-渗透率	0.880	1.967	8.515	0.2530	0.1516	0.0023
4-经营效率	0.195	0.434	0.955	0.6656	0.5886	0.5140
5-盈利改善	0.248	0.664	0.425	0.1806	0.0513	0.5032
6-财务稳健性	-0.342	-0.758	-0.757	0.2821	0.1870	0.4882
7-业绩对比	0.016	0.427	-0.318	0.9439	0.3093	0.6918
8-股东行为	0.412	-0.259	0.593	0.6457	0.8708	0.8356
9-政策驱动	-0.482	4.554	10.075	0.4706	0.0033	0.0003
10-市场驱动	-0.085	0.262	3.407	0.8377	0.7296	0.0169

资料来源：Wind 资讯，慧博投研，华安证券研究所

4.3.2 篇幅：重点说明 vs 轻描淡写的信号效应

进一步考察情绪类别在摘要中所占篇幅比例是否具有预测性，回归模型如下：

$$CAR5/CAR20/CAR60 = \beta_0 + \beta_1 * \text{归一化篇幅占比} + \beta_2 * \text{行业虚拟变量} + \epsilon$$

核心结论如下：

- (1) 利空信息的篇幅占比与未来收益整体相关性不显著，说明其市场定价中的作用不够稳定，可能受到舆情环境或叠加事件影响。
- (2) 对于利好信息，部分类别展现出一定的正向关系，尤其是**3-渗透率**，**9-政策驱动**，**2-技术创新**，**1-战略布局**。这表明，当这些因素在研报摘要中被着重描述时，市场更倾向于将其解读为正向信号，从而提升个股未来收益概率。
- (3) 值得注意的是，对于**5-盈利改善**与**6-财务稳健性**这两个基本面强相关类别，在短期（5日、20日）窗口中，利好的篇幅占比与未来收益甚至呈**负相关**。可能原因在于：相关研报多为业绩点评类，天然会占据较大篇幅；高频出现的“基本面陈述”未必具备稀缺性或超额信息含量，篇幅反而不能作为正向信号指标。

图表 33 篇幅占比与 CAR 的回归结果（情绪值为+1）

利好类别	回归系数			系数显著性 (P-Value)		
	CAR5	CAR20	CAR60	CAR5	CAR20	CAR60
1-战略布局	0.622	0.188	1.019	0.0001	0.5100	0.0462
2-技术创新	0.950	1.595	3.392	0.0000	0.0000	0.0000
3-渗透率	1.276	2.163	3.894	0.0000	0.0000	0.0000
4-经营效率	0.224	-0.313	2.239	0.3864	0.5068	0.0093
5-盈利改善	0.049	-0.548	3.114	0.7423	0.0412	0.0000
6-财务稳健性	-0.025	-0.551	1.298	0.9162	0.1963	0.1083
7-业绩对比	2.168	1.184	2.897	0.0000	0.1207	0.0472
8-股东行为	0.442	0.156	3.908	0.1046	0.7542	0.0000
9-政策驱动	1.115	2.602	3.454	0.0236	0.0034	0.0237
10-市场驱动	0.469	0.007	2.499	0.2033	0.9921	0.0475

资料来源：Wind 资讯，慧博投研，华安证券研究所

图表 34 篇幅占比与 CAR 的回归结果（情绪值为-1）

利好类别	回归系数			系数显著性 (P-Value)		
	CAR5	CAR20	CAR60	CAR5	CAR20	CAR60
1-战略布局	-0.201	-11.889	-11.403	0.9508	0.1370	0.3076
2-技术创新	-1.893	-6.148	-15.943	0.7455	0.5909	0.4530
3-渗透率	2.565	4.268	12.125	0.2034	0.2348	0.1148
4-经营效率	-1.216	2.781	11.182	0.2359	0.1274	0.0007
5-盈利改善	-0.319	0.324	3.421	0.1310	0.3968	0.0000
6-财务稳健性	-0.189	0.518	7.679	0.7410	0.6152	0.0001
7-业绩对比	-0.750	-1.926	-1.648	0.2014	0.0720	0.4382
8-股东行为	1.102	-3.034	6.237	0.6122	0.4300	0.3671
9-政策驱动	-1.433	3.189	9.403	0.3539	0.3757	0.1614
10-市场驱动	-0.192	-0.693	0.132	0.8109	0.6378	0.9639

资料来源：Wind 资讯，慧博投研，华安证券研究所

5 研报摘要综合因子的构建与评估

在前文中，我们从情绪类别方向、摘要顺序、情绪密度及篇幅占比等多个维度探讨了研报内容结构与未来收益之间的关系。接下来，我们进一步整合上述信息，构建研报层面和个股层面的综合评分指标，以衡量其对未来收益的预测能力，并探索实际的应用价值。

5.1 报告维度的综合评分算法

首先从报告维度计算报告本身的得分，我们尝试了多种评分方法，并分析多种方式对未来超额收益（CAR5/CAR20/CAR60）的回归分析结果，主要比较了它们在回归系数、显著性、模型解释力（R²）和 Rank IC（日度横截面预测力）四个维度的表现。

5.1.1 简单加权及集中度调整

我们首先尝试采用分类情绪值的简单累加或均值计算方式，对研报整体的正负

信息倾向进行评分。

其中,直接累加各类别情绪值得到的总得分为 $score_sum$ 。这一方法的问题在于,大多数研报仅覆盖 3-4 类情绪,导致得分集中度高,且得分高度依赖于“提及类别的数量”,即使各类内容仅为轻描淡写,类别越多,得分也越高。相比之下,情绪均值得分($score_mean$)试图剔除类别数量影响,但该方法同样存在两个缺陷:(1)结果集中度过高,不利于区分不同摘要;(2)完全忽视了利好类别数量的价值。例如,摘要 A 仅涉及 1 个情绪为正向的类别,摘要 B 涉及 3 个情绪为正向的类别,两者的得分均为 1。

$$score_sum = \sum_{k=1}^{10} cat_sentiment_k$$

$$score_mean = \sum_{k=1}^{10} cat_sentiment_k / cat_count$$

从实际视角看,类别数量不仅代表信息广度,也体现某种“风险分散”效应,单纯取均值的方法容易高估那些篇幅集中于少数利好类别的报告,低估那些结构均衡、信息覆盖更广的报告。若一份摘要主要依赖于某一强利好(如补贴政策),其得分虽高,但信息高度集中,一旦该利好退坡,企业业绩可能出现剧烈波动。而覆盖多个利好类别的摘要,则通常意味着企业在财务、战略、市场等多个维度均表现积极,具备更强的稳健性与抗风险能力,因此理应获得更高的评价。为反映这种“多样化”特征的重要性,我们引入集中度惩罚机制,即在均值基础上加入赫芬达尔指数(HHI)调整,构建 $score_mean_hhi$ 。该机制通过衡量摘要中各类别的篇幅占比(cat_len_k)来计算集中度:

$$HHI = \sum_{k=1}^{10} cat_len_k^2$$

$$score_mean_hhi = score_mean * 1/\sqrt{0.01 + HHI}$$

其中, cat_len_k 表示第 k 类在摘要中的篇幅占比,取值范围为 0-1。HHI 越高,表示信息越集中,得分将被更多惩罚,从而鼓励多类别、低集中度的信息结构。

图表 35 的结果显示,两者在短期(CAR5)和中期(CAR20)均具显著正向预测作用,尤其是经过 HHI 调整后 $score_mean_hhi$ 的回归系数更大、P 值更小,预测力更强;而在长期(CAR60)则呈现负相关,且不具显著性。此外,从 CAR20 的 Rank IC 看, $score_mean_hhi$ 的 RankIC 也优于 $score_sum$ 。总体而言, $score_mean_hhi$ 在收益预测中的效果更优。

图表 35 $score_sum$ 和 $score_mean_hhi$ 与 CAR 的回归结果

	回归系数			系数P-Value			模型R2			日 Rank IC
	CAR5	CAR20	CAR60	CAR5	CAR20	CAR60	CAR5	CAR20	CAR60	CAR20
$score_sum$	0.10	0.04	-0.02	0.00	0.01	0.49	0.40%	0.67%	1.11%	2.57%
$score_mean_hhi$	0.28	0.08	-0.08	0.00	0.00	0.13	0.46%	0.76%	1.35%	2.94%

资料来源: Wind 资讯, 慧博投研, 华安证券研究所整理

5.1.2 篇幅占比加权

为了更准确反映研报摘要中信息的表达强度,我们引入各类别在摘要中所占的篇幅比例,将情绪值按篇幅加权。这样可以有效区分“重点阐述”与“轻描淡写”的差异。例如,两份报告都提及“技术创新”,但若一份用 10%的篇幅深入展开,另一份仅一句带过,前者的加权得分更高,更贴近信息的实际重要性。

$$score_by_len = \sum_{k=1}^{10} cat_sentiment_k * cat_len_k$$

同样的，仅采用篇幅占比加权的方法 ($score_by_len$)，仍可能高估内容集中于少数类别的摘要，从而低估覆盖更广、内容更丰富的报告。因此，我们同样在此基础上引入**信息集中度惩罚机制**，通过以下两种方式强化“信息分散”在得分中的权重：

(1) **HHI 惩罚调整 ($score_by_len_hhi$)**

$$score_by_len_hhi = score_by_len * 1/\sqrt{0.01 + HHI}$$

(2) **幂函数平滑调整**：在前述分析中我们也观察到，篇幅本身在部分类别上并不总是有效信息，甚至可能引入噪声。因此，我们在保留篇幅权重的同时，对其影响力进行“弱化处理”，以避免其对整体评分产生过强主导作用。我们采用幂函数压缩方式对篇幅占比进行调整，通过这种非线性处理方式，即使某个类别篇幅占比较高，其影响力也不会线性上升，有效防止“篇幅主导效应”的出现；同时，中等篇幅的类别也能获得合理的权重。

$$score_by_len_alpha = \sum_{k=1}^{10} (cat_sentiment_k * cat_len_k^\alpha), \alpha = 0.2$$

分析表明，引入集中度调整后， $score_by_len_hhi$ 和 $score_by_len_alpha$ 相比原始指标在预测排序效果 (Rank IC) 和解释力 (R^2) 上均有所提升，尽管回归系数略低，但统计显著性更强，表现更稳健，适合作为更有效的收益预测因子。

图表 36 $score_by_len$ 与 CAR 的回归结果

	回归系数			系数P-Value			模型R2			日 Rank IC
	CAR5	CAR20	CAR60	CAR5	CAR20	CAR60	CAR5	CAR20	CAR60	CAR20
score_by_len	0.51	0.21	0.08	0.00	0.00	0.38	0.44%	0.55%	1.11%	2.22%
score_by_len_hhi	0.27	0.07	-0.12	0.00	0.01	0.01	0.46%	0.75%	1.34%	2.90%
score_by_len_alpha	0.16	0.04	-0.11	0.00	0.02	0.00	0.44%	0.75%	1.34%	2.85%

资料来源：Wind 资讯，慧博投研，华安证券研究所整理

5.1.3 类别重要性加权

我们前述的加权方法隐含一个前提假设：所有类别的影响力是相同的。然而，从我们此前各个类别的“情绪方向与未来收益关系”的实证分析可以看出，不同类别对收益的预测能力存在显著差异：某些类别表现出更强的相关性，有些则几乎无预测作用，甚至呈现反向关系。因此，有必要引入反映类别重要性的权重体系，从而提高综合得分的含义。

1. 类别权重的来源

我们以 CAR 为目标变量，构建如下多元回归模型，对各类别情绪的收益贡献进行量化，以 CAR20 为例：

$$CAR20 = \beta_0 + \beta_{1,k} * \sum_{k=1}^{10} cat_sentiment_k + \beta_{2,j} * \sum_{j=1}^{30} industry_dummy_j + \epsilon$$

其中， $cat_sentiment_k$ 表示第 k 类别的情绪值（若报告未涉及该类别则为 0）， $industry_dummy_j$ 控制行业异质性。通过该模型，我们可以获得每一类情绪的回归系数 $\beta_{1,k}$ 及其统计显著性，进而用作类别的“预测贡献度”指标。

2. 类别重要性权重方案

理论上，回归系数为正且统计显著，代表该类别正向情绪具有正向收益预测力，其系数越大越可靠。因此，我们基于不同策略设计了五种确定类别权重 (cat_w) 的

方案。

图表 37 类别重要性权重的计算方案

权重	权重名	说明
cat_w1	回归系数加权	直接使用基于 CAR20 回归所得到的每个类别的回归系数 $\beta_{-}(1,k)$ 作为其对应权重 cat_w_k。
cat_w2	系数顺序加权	根据 CAR20 的回归系数的大小进行降序排名，并转换为比例权重： $w3 = (10 - rank + 1) / 10$ ，即系数越大，权重越高。
cat_w3	基于 CAR20 与 CAR60 的系数有效性综合赋权	建立“打分”机制判断类别的重要性： i. CAR20：若某类别在 CAR20 回归中系数为正且在 10% 显著性水平内，得 2 分；若正但不显著，得 1 分； ii. CAR60：若在 CAR60 回归中也为正且显著，则再加 1 分。 最后将所有类别得分进行归一化，得到综合权重。
权重的后处理与边界控制		
考虑到极端情况与实务可行性，我们引入如下两项后处理机制，最后，再对类别权重进行归一		
	1) 极端情况调整：	有效类别太少再调整，主要针对仅根据 CAR20 的系数进行赋权的情况。具体而言，若有效权重（即 $w > 0$ ）的类别数 < 5 ，为避免模型过度依赖少数类别，将引入 CAR60 回归结果补充权重： i. 若 CAR60 中该类别系数为正且显著，则赋予当前最小有效权重； ii. 若正但不显著，则赋为当前最小有效权重的一半。
	2) 类别权重上限限制：	为避免单一类别权重过高而造成“结构性倾斜”，设定单类权重上限为 0.3。对超过上限的类别进行裁剪，并将剩余权重按比例重新分配，最后归一化处理。

资料来源：Wind 资讯，华安证券研究所整理

3. 权重计算频率与实际回归机制

在实际操作中，我们采用月频滚动回归法动态估计类别权重。具体而言，每月以过去 12 个月内发布的报告为样本，计算各类别的回归系数与显著性，更新最新权重 cat_w_k。以此为基础，我们构造多个版本的综合评分指标：

$$score_by_cat_w\{i\} = \sum_{k=1}^{10} (cat_sentiment_k * cat_w_k\{i\}), i = 1, 2, 3$$

$$score_by_cat_w\{i\}_hhi = score_by_cat_w\{i\} * 1/\sqrt{0.01 + HHI}$$

除了以上基于权重加总的评分方案外，我们还构建了直接的收益预测模型，将模型预测值作为报告的“潜在得分”，记为 score_by_cat_pred，代表综合情绪对 CAR20 的预测值。

$$\widehat{CAR20} = \beta_0 + \beta_{1,k} * \sum_{k=1}^{10} cat_sentiment_k + \beta_{2,j} * \sum_{j=1}^{30} industry_dummy_j + \epsilon$$

结果表明：

- (1) 所有打分方式的回归系数均显著为正，报告得分与未来收益正相关。各打分因子（包括加权得分 w1/w2/w3 及其集中度修正版本_hhi）在 CAR5、CAR20、CAR60 上都具有显著正系数，P 值为 0，表明利好情绪分数与未来超额收益显著正相关。
- (2) 加入集中度修正(_hhi)对效果边际影响较小。加入_hhi 的版本系数略低，模型 R² 变化不大，说明情绪分布集中度对收益预测的增强作用有限，更多体现为稳健性调整。其本质是因为 score_by_cat 本身的算法便已经考虑了类别数量因素，涉及类别多的报告倾向于得到更高分，因此再度进行 HHI 修正的意义并不大。
- (3) w3（根据回归系数和显著性设置分级权重）表现最优。在短中期收益预测（CAR5、CAR20）和横截面排序（Rank IC）方面表现最佳；尽管长期预测（CAR60）略低于 w1，但差距不大；结合回归系数、R²、Rank IC 三

个维度，w3 是综合表现最优的加权方式。

- (4) *score_by_cat_pred* 的 Rank IC 最高，但回归解释力偏弱。*score_by_cat_pred*在 CAR20 上的 Rank IC 高达 4.00%，显著优于其他因子 (2.4~2.8%)；但它的回归系数仅为 0.15~0.59，模型 R²显著低于其他方法，说明其在截面排序上的指示性强，但解释未来收益幅度有限。

图表 38 *score_by_cat*与 CAR 的回归结果

	回归系数			系数P-Value			模型R2			日 Rank IC
	CAR5	CAR20	CAR60	CAR5	CAR20	CAR60	CAR5	CAR20	CAR60	CAR20
<i>score_by_cat_w1</i>	1.35	1.00	2.51	0.00	0.00	0.00	0.54%	0.79%	1.45%	2.64%
<i>score_by_cat_w1_hhi</i>	0.62	0.43	1.00	0.00	0.00	0.00	0.51%	0.78%	1.43%	2.82%
<i>score_by_cat_w2</i>	1.35	0.97	2.02	0.00	0.00	0.00	0.51%	0.78%	1.41%	2.40%
<i>score_by_cat_w2_hhi</i>	0.59	0.38	0.63	0.00	0.00	0.00	0.48%	0.78%	1.41%	2.46%
<i>score_by_cat_w3</i>	1.42	1.16	2.24	0.00	0.00	0.00	0.53%	0.80%	1.35%	3.29%
<i>score_by_cat_w3_hhi</i>	0.64	0.48	0.81	0.00	0.00	0.00	0.51%	0.78%	1.40%	3.23%
<i>score_by_cat_pred</i>	0.15	0.22	0.59	0.00	0.00	0.00	0.36%	0.31%	1.34%	4.00%

资料来源：Wind 资讯，慧博投研，华安证券研究所整理

5.1.4 篇幅占比&类别重要性加权

在前文中，我们已经分别探讨了基于篇幅占比加权 (*score_by_len*) 与类别重要性加权 (*score_by_cat*) 的得分方法，这两种方法各有优劣。综合考虑摘要的篇幅密度与类别的重要性，提出了篇幅占比与类别重要性加权的结合方法。

$$score_by_LenCat\{i\} = \sum_{k=1}^{10} (cat_sentiment_k * cat_len_k * cat_w_k\{i\}), i = 1,2,3$$

为了避免得分受到单一类别篇幅占比的过度主导，我们再次引入了赫芬达尔指数 (HHI) 和幂函数调整，计算集中度调整后的篇幅占比&类别重要性加权得分：

$$score_by_LenCat\{i\}_hhi = score_by_LenCat\{i\} * 1/\sqrt{0.01 + HHI}, i = 1,2,3$$

$$score_by_LenCat\{i\}_alpha = \sum_{k=1}^{10} (cat_sentiment_k * cat_len_k^2 * cat_w_k\{i\}), i = 1,2,3$$

结果表明：

- (1) 引入篇幅加权显著提升回归系数。原始 *score_by_LenCat* 系列 (无后缀) 在所有时间窗口下的回归系数最高，说明篇幅占比 × 情绪值 × 类别权重三重加权后，打分对未来收益有强解释力。
- (2) 引入 HHI (集中度) 保持预测力，部分优化排序能力。*_hhi* 版本的回归系数略低，但模型 R² 与 Rank IC 与原始版本非常接近，表明情绪集中度的修正不会损害预测力，可作为稳健性增强措施。
- (3) 引入 α (弱化篇幅权重) 权衡解释力与排序能力。*_alpha* 版本普遍具有更高的日 Rank IC，但回归系数显著降低；说明适度削弱篇幅权重虽会降低幅度预测力，但有助于增强截面选股能力。
- (4) *LenCat3* 系列表现最强。在三个版本中，w3 加权方式在 R² 和 Rank IC 上保持领先，无论是基础版、*_hhi* 还是 *_alpha*，表明其在收益解释与排序预测上的综合能力最强。

图表 39 score_by_LenCat与 CAR 的回归结果

	回归系数			系数P-Value			模型R2			日 Rank IC
	CAR5	CAR20	CAR60	CAR5	CAR20	CAR60	CAR5	CAR20	CAR60	CAR20
score_by_LenCat1	5.53	3.84	11.55	0.00	0.00	0.00	0.59%	0.79%	1.49%	2.32%
score_by_LenCat1_alpha	1.88	1.28	3.42	0.00	0.00	0.00	0.56%	0.79%	1.46%	2.63%
score_by_LenCat1_hhi	3.12	2.17	6.30	0.00	0.00	0.00	0.59%	0.79%	1.48%	2.63%
score_by_LenCat2	5.94	3.91	10.45	0.00	0.00	0.00	0.58%	0.79%	1.46%	1.90%
score_by_LenCat2_alpha	1.92	1.24	2.67	0.00	0.00	0.00	0.54%	0.78%	1.44%	2.31%
score_by_LenCat2_hhi	3.29	2.15	5.37	0.00	0.00	0.00	0.57%	0.79%	1.45%	2.28%
score_by_LenCat3	6.07	4.34	10.13	0.00	0.00	0.00	0.60%	0.80%	1.44%	3.16%
score_by_LenCat3_alpha	2.07	1.53	3.01	0.00	0.00	0.00	0.57%	0.79%	1.43%	3.35%
score_by_LenCat3_hhi	3.44	2.52	5.38	0.00	0.00	0.00	0.60%	0.80%	1.44%	3.50%

资料来源: Wind 资讯, 慧博投研, 华安证券研究所整理

5.2 个股层面的评分指标构建

从报告评分与未来收益的相关性来看, 研报评分在研报发布后 20 日、40 日具有显著的收益解释力, 体现出较强的事件驱动特征。那么, 从截面选股角度是否同样具备选股能力? 为此, 我们进一步将报告评分延伸至个股层面, 尝试将其转化为可实际应用的因子。基于所有研报样本, 我们构建了以下两种算法的评分因子, 用于评估其在截面选股中的表现。

- (1) **等权评分因子 (avgRecord)**: 滚动窗口期内所有有效记录的简单等权, 并非先计算日均评分再日间等权。如近 3 日有 10 个有效的评分记录, 则直接计算 10 个记录的均值, 反映了分析师的整体看法, 预期为**正向因子**, 即高的均匀评分与较高的未来收益有关。
- (2) **指数衰减加权因子 (ewaRecord)**: 根据时间对评分进行加权, 使最近的评分得到更大的权重。这种加权方式能够确保因子更加关注近期的市场情绪或分析师的看法, 认为最近的评价更能反映股票的当前或未来表现。理论上, 时间衰减评分因子也是**正向因子**。

$$w_{i_num} = 2^{\frac{i-N-1}{H}}, i = 1 \dots N, H = N/2$$

$$w_{i_den} = \sum w_{i_num}$$

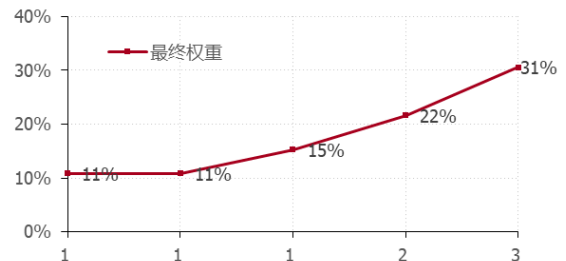
$$\text{最终权重 } w_i = w_{i_num} / w_{i_den}$$

图表 40 指数衰减加权的权重计算举例

	日期标记(i)	报告得分	初始权重	最终权重
2022/3/20	1	8.5	25%	11%
2022/3/20	1	8.6	25%	11%
2022/3/20	1	9	25%	11%
2022/3/24	2	9	35%	15%
2022/3/25	3	8	50%	22%
2022/3/26	4	9.5	71%	31%
合计			231%	100%

资料来源: wind 资讯, 华安证券研究所

图表 41 指数衰减加权的权重示意图



资料来源: wind 资讯, 华安证券研究所

基于此, 我们在每个月末观察日, 根据过去 1、2 个月的报告得分记录计算个股的评分, 记作_1m 和_2m。近 1 个月窗口期的因子覆盖度为 32%, 近 2 个月窗口期的因子覆盖度为 42%。

图表 42 个股评分因子及逻辑

报告得分算法	报告因子	个股评分算法	评分窗口期	个股因子举例
简单加权	score_mean_hhi	等权加权 (_avgRecord) 指数衰减加权 (_ewaRecord)	1. 近1、2个月的有效评分记录 (_1m, _2m) 2. 缺失值填充逻辑: 缺失值用近N个月内的最新有效值进行填充	score_mean_hhi_ewa_1m
篇幅占比加权	score_by_len_hhi			score_by_len_hhi_avg_2m
	score_by_len_alpha			score_by_len_alpha_avg_1m
类别重要性加权	score_by_cat_w3			score_by_cat_w3_avg_1m
	score_by_cat_w3_hhi	score_by_cat_w3_hhi_avg_2m		
篇幅占比&类别重要性加权	score_by_LenCat3_hhi	score_by_LenCat3_hhi_avg_1m		
	score_by_LenCat3_alpha	score_by_LenCat3_alpha_ewa_1m		

资料来源: 华安证券研究所整理

因子检测参数如下:

- 测试区间: 2019.12.31 – 2025.05.30
- 股票池: 沪深交易所股票, 剔除上市不足 90 天、ST、*ST、停牌、ST 或 *ST 摘帽不足 30 天的个股
- 因子处理: 3 倍标准差去极值、行业中位数填充、zs 标准化、行业市值中性化 (中信一级), 仅保留有因子数据的部分
- 交易处理: 考虑涨跌停、停牌等行为导致的交易可行性
- 组数: 5 分组, 基准: 中证 800 (000906.SH)
- 调仓频率: 月频

1. 简单加权及集中度调整

图表 43 展示了均值 HHI 调整因子 (score_mean_hhi) 在不同加权方式和窗口期下的表现。整体来看, 1 个月窗口期的效果优于 2 个月, 指数衰减加权优于等权加权, 反映出数据越接近当前、预测力越强。

综合因子的覆盖度与有效性, 我们采用指数衰减加权的方式, 将 1m 和 2m 因子的权重分别设为 2/3 和 1/3, 构建复合因子 score_mean_hhi。该因子的 Rank IC 为 1.72%, IC 胜率 67.7%, 多头年化收益为 11.6%, 相对中证 800 的年化超额为 12.3%, 具备一定的收益捕捉能力。不过, 由于缺乏有效的空头识别, 整体多空策略表现较弱。

图表 43 score_mean 因子表现

因子	t值	Rank IC	ICIR	IC胜率	多头年化	多头年化超额	多头胜率	多头夏普	多头最大回撤	选择	权重
score_mean_hhi_avg_1m	2.11	1.74%	0.95	67.7%	11.9%	12.6%	67.7%	0.42	-30.9%		
score_mean_hhi_avg_2m	2.64	1.51%	0.73	67.7%	10.7%	11.4%	63.1%	0.37	-31.0%		
score_mean_hhi_ewa_1m	2.10	1.85%	1.02	67.7%	12.9%	13.6%	67.7%	0.47	-30.9%	✓	2/3
score_mean_hhi_ewa_2m	2.61	1.54%	0.76	66.2%	11.7%	12.4%	64.6%	0.41	-30.2%	✓	1/3
score_mean_hhi	2.39	1.72%	0.88	67.7%	11.6%	12.3%	67.7%	0.41	-30.9%		

资料来源: Wind 资讯, 慧博投研, 华安证券研究所整理

2. 篇幅占比加权

将基于篇幅长度加权的因子进行 HHI 调整 (_hhi) 与篇幅弱化调整 (_alpha) 后, 两类因子的整体表现相近, RankIC 和收益水平与 score_mean_hhi 因子相比并无显著提升, 仅有多头胜率上略有优势。

图表 44 *score_by_len* 因子表现

因子	t值	Rank IC	ICIR	IC胜率	多头 年化	多头年 化超额	多头 胜率	多头 夏普	多头最 大回撤	选 择	权重
<i>score_by_len_hhi_avg_1m</i>	2.12	1.63%	0.87	67.7%	12.4%	13.1%	66.2%	0.45	-30.8%		
<i>score_by_len_hhi_avg_2m</i>	2.63	1.39%	0.67	66.2%	10.9%	11.5%	63.1%	0.38	-30.8%		
<i>score_by_len_hhi_ewa_1m</i>	2.12	1.72%	0.94	69.2%	12.3%	13.0%	63.1%	0.44	-30.6%	✓	2/3
<i>score_by_len_hhi_ewa_2m</i>	2.60	1.39%	0.68	64.6%	11.8%	12.5%	64.6%	0.42	-30.7%	✓	1/3
<i>score_by_len_hhi</i>	2.40	1.62%	0.82	66.2%	12.0%	12.7%	69.2%	0.43	-30.8%		
<i>score_by_len_alpha_avg_1m</i>	2.08	1.64%	0.9	69.2%	13.0%	13.7%	66.2%	0.47	-31.3%		
<i>score_by_len_alpha_avg_2m</i>	2.59	1.43%	0.69	67.7%	11.5%	12.2%	64.6%	0.4	-31.0%		
<i>score_by_len_alpha_ewa_1m</i>	2.07	1.72%	0.95	69.2%	13.1%	13.9%	69.2%	0.47	-31.5%	✓	2/3
<i>score_by_len_alpha_ewa_2m</i>	2.57	1.45%	0.71	67.7%	11.1%	11.8%	63.1%	0.39	-31.1%	✓	1/3
<i>score_by_len_alpha</i>	2.35	1.57%	0.81	67.7%	11.4%	12.1%	66.2%	0.4	-32.4%		

资料来源：Wind 资讯，慧博投研，华安证券研究所整理

3. 类别重要性加权

采用类别重要性加权的因子相较于前边的篇幅加权因子 (*score_by_len*)，在多头年化收益和多头胜率上表现更优，而在 RankIC 表现上差异不大，且 RankIC 胜率略逊一筹。

- 类别加权因子 *score_by_cat_w3* 和 *score_by_cat_w3_hhi* 的多头年化收益分别为 13.1% 和 13.3%，相对中证 800 的年化超额为 13.8% 和 14.0%。
- 篇幅加权因子 *score_by_len_hhi* 和 *score_by_len_alpha* 的 RankIC 胜率分别为 66.2% 和 67.7%，高于类别加权因子的 61.5% 和 63.1%。

图表 45 *score_by_cat* 因子表现

因子	t值	Rank IC	ICIR	IC胜率	多头 年化	多头年 化超额	多头 胜率	多头 夏普	多头最 大回撤	选 择	权重
<i>score_by_cat_w3_avg_1m</i>	1.93	1.85%	1.01	63.1%	13.1%	13.8%	61.5%	0.47	-32.8%		
<i>score_by_cat_w3_avg_2m</i>	2.31	1.54%	0.76	63.1%	12.2%	12.9%	60.0%	0.43	-31.8%		
<i>score_by_cat_w3_ewa_1m</i>	1.90	1.87%	1.04	64.6%	13.0%	13.7%	64.6%	0.46	-33.1%	✓	2/3
<i>score_by_cat_w3_ewa_2m</i>	2.30	1.55%	0.79	63.1%	11.6%	12.3%	61.5%	0.4	-32.4%	✓	1/3
<i>score_by_cat_w3</i>	2.16	1.63%	0.88	61.5%	13.1%	13.8%	61.5%	0.47	-32.6%		
<i>score_by_cat_w3_hhi_avg_1m</i>	1.86	1.80%	1.01	64.6%	14.8%	15.5%	61.5%	0.54	-33.2%		
<i>score_by_cat_w3_hhi_avg_2m</i>	2.25	1.49%	0.75	63.1%	12.6%	13.3%	60.0%	0.44	-32.1%		
<i>score_by_cat_w3_hhi_ewa_1m</i>	1.83	1.83%	1.04	63.1%	14.2%	14.9%	60.0%	0.51	-33.5%	✓	2/3
<i>score_by_cat_w3_hhi_ewa_2m</i>	2.23	1.52%	0.79	63.1%	12.0%	12.7%	60.0%	0.41	-33.3%	✓	1/3
<i>score_by_cat_w3_hhi</i>	2.09	1.58%	0.89	63.1%	13.3%	14.0%	63.1%	0.47	-32.7%		

资料来源：Wind 资讯，慧博投研，华安证券研究所整理

4. 篇幅占比&类别重要性加权

综合篇幅占比与类别重要性加权后的调整因子 (*score_by_LenCat3* 系列) 在 RankIC、多头收益和胜率上整体优于单一维度加权的因子，但在 IC 胜率上略低。

图表 46 *score_by_LenCat* 因子表现

因子	t值	Rank IC	ICIR	IC胜率	多头 年化	多头年 化超额	多头 胜率	多头 夏普	多头最 大回撤	选 择	权重
<i>score_by_LenCat3_hhi_avg_1m</i>	1.96	1.76%	0.94	64.6%	12.5%	13.2%	61.5%	0.44	-31.9%		
<i>score_by_LenCat3_hhi_avg_2m</i>	2.36	1.60%	0.8	64.6%	11.5%	12.2%	58.5%	0.4	-30.7%		
<i>score_by_LenCat3_hhi_ewa_1m</i>	1.93	1.77%	0.96	63.1%	13.1%	13.8%	58.5%	0.47	-31.7%	✓	2/3
<i>score_by_LenCat3_hhi_ewa_2m</i>	2.35	1.56%	0.79	63.1%	12.1%	12.8%	63.1%	0.43	-30.7%	✓	1/3
<i>score_by_LenCat3_hhi</i>	2.17	1.65%	0.87	63.1%	12.1%	12.9%	55.4%	0.43	-31.8%		
<i>score_by_LenCat3_alpha_avg_1m</i>	1.96	1.86%	1.00	63.1%	12.8%	13.5%	60.0%	0.45	-32.7%		
<i>score_by_LenCat3_alpha_avg_2m</i>	2.34	1.57%	0.78	63.1%	11.8%	12.5%	60.0%	0.41	-31.3%		
<i>score_by_LenCat3_alpha_ewa_1m</i>	1.93	1.88%	1.03	64.6%	13.6%	14.3%	60.0%	0.49	-32.9%	✓	2/3
<i>score_by_LenCat3_alpha_ewa_2m</i>	2.33	1.57%	0.79	61.5%	11.4%	12.1%	63.1%	0.4	-32.5%	✓	1/3
<i>score_by_LenCat3_alpha</i>	2.18	1.66%	0.89	61.5%	13.5%	14.2%	58.5%	0.48	-32.1%		

资料来源：Wind 资讯，慧博投研，华安证券研究所整理

5. 基于大模型的研报摘要分类的综合评分因子

接下来，我们将上述各类小类因子等权合成为四类核心因子：均值 HHI 调整因子 *score_mean_hhi*、篇幅加权因子 *score_by_len*、类别重要性加权因子 *score_by_cat*，以及篇幅与类别重要性双重加权的因子 *score_by_LenCat*。其中，*score_by_cat* 和 *score_by_LenCat* 在多头年化收益和超额收益上表现最佳，相对中证 800 的年化超额分别为 13.9% 和 13.7%；而 *score_mean_hhi* 则在 RankIC (1.72%) 和 IC 胜率 (67.7%) 上更具优势。

图表 47 基于研报摘要分类的合成因子的表现

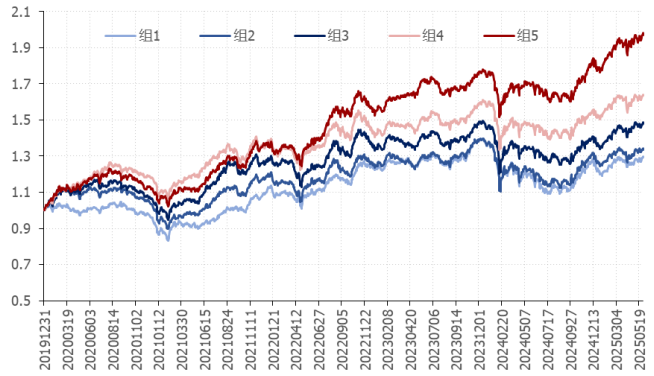
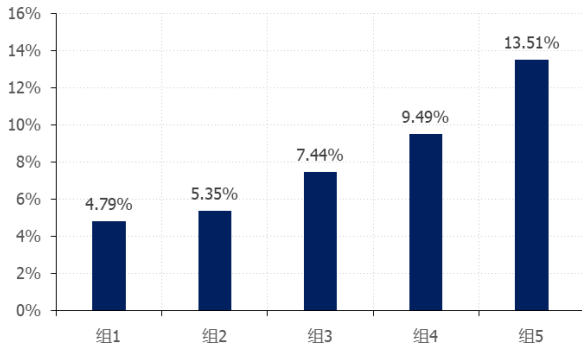
因子	t值	Rank IC	ICIR	IC胜率	多头年化	多头年化超额	多头胜率	多头夏普	多头最大回撤
<i>score_mean_hhi</i>	2.39	1.72%	0.88	67.7%	11.6%	12.3%	67.7%	0.41	-30.9%
<i>score_by_len_hhi</i>	2.40	1.62%	0.82	66.2%	12.0%	12.7%	69.2%	0.43	-30.8%
<i>score_by_len_alpha</i>	2.35	1.57%	0.81	67.7%	11.4%	12.1%	66.2%	0.40	-32.4%
<i>score_by_len</i>	2.39	1.58%	0.80	67.7%	11.7%	12.4%	64.6%	0.41	-31.9%
<i>score_by_cat_w3</i>	2.16	1.63%	0.88	61.5%	13.1%	13.8%	61.5%	0.47	-32.6%
<i>score_by_cat_w3_hhi</i>	2.09	1.58%	0.89	63.1%	13.3%	14.0%	63.1%	0.47	-32.7%
<i>score_by_cat</i>	2.13	1.61%	0.89	63.1%	13.2%	13.9%	66.2%	0.47	-33.1%
<i>score_by_LenCat3_hhi</i>	2.17	1.65%	0.87	63.1%	12.1%	12.9%	55.4%	0.43	-31.8%
<i>score_by_LenCat3_alpha</i>	2.18	1.66%	0.89	61.5%	13.5%	14.2%	58.5%	0.48	-32.1%
<i>score_by_LenCat</i>	2.19	1.67%	0.89	61.5%	13.0%	13.7%	58.5%	0.46	-31.9%
<i>score_report_llm</i>	2.37	1.77%	0.91	66.2%	12.8%	13.5%	64.6%	0.46	-31.5%

资料来源：Wind 资讯，慧博投研，华安证券研究所整理

在此基础上，我们将上述四类核心因子进一步等权合成为最终的综合评分因子 *score_report_llm*，该因子在各项指标上表现相对更均衡，有一定的选股能力，多头分年度表现稳定且超额回撤控制良好。除此之外，合成因子和传统因子基本没有相关性，可作为另类因子。

- 整体指标来看，RankIC 为 1.77%、IC 胜率 66.2%、ICIR 为 0.91，具备一定的横截面选股能力；多头年化收益为 12.8%、年化超额收益为 13.5%。
- 分组方面，五组收益严格单调，组间收益差异明确，尽管组内表现存在一定波动，整体结构具备辨识度。
- 多头年度表现方面，自 2020 至 2025 年 5 月，多头组合相对中证 800 的年胜率为 100%；今年的调仓胜率为 100%，超额收益为 10.4%。此外，相对五分组合等权组合，各年度超额最大回撤均控制在 4% 以内，2025 年更仅为 1.2%。
- *score_report_llm* 因子与传统因子的相关性低，与一致预期调整因子的相关性最大，也仅有 0.185，可认为是一个全新的另类因子。

图表 48 score_report_llm的分组年化超额 (中证 800) 图表 49 score_report_llm的分组超额净值 (中证 800)



资料来源: Wind 资讯, 慧博投研, 华安证券研究所

资料来源: Wind 资讯, 慧博投研, 华安证券研究所

图表 50 score_report_llm因子的多头分年度表现

	收益率	中证 800	超额收益	最大回撤	超额最大回撤	信息比率	调仓胜率-中证800	调仓超额均值-中证800	调仓盈亏比
2020	36.5%	25.8%	10.7%	-15.5%	-11.5%	0.9	54.5%	0.24%	4.5
2021	25.4%	-0.8%	26.2%	-8.5%	-7.1%	2.5	63.6%	2.33%	1.2
2022	-10.8%	-21.3%	10.5%	-31.3%	-9.8%	1.1	54.5%	1.55%	1.2
2023	1.2%	-10.4%	11.6%	-15.5%	-6.4%	1.7	90.9%	0.99%	0.6
2024	12.3%	12.2%	0.1%	-21.9%	-14.4%	0.0	54.5%	0.87%	1.8
20250530	8.4%	-2.0%	10.4%	-14.7%	-4.7%	2.7	100.0%	1.64%	3.0
整体	12.8%	-0.7%	13.5%	-31.5%	-17.0%	1.2	64.6%	1.08%	1.5

资料来源: Wind 资讯, 慧博投研, 华安证券研究所整理

图表 51 score_report_llm因子的相关性

	分析师调整因子	一致预期调整因子	SUE 因子	成长因子
score_report_llm	0.143	0.185	0.123	0.099

资料来源: Wind 资讯, 慧博投研, 华安证券研究所整理

6 总结

本文基于 DeepSeek 大模型, 探索将情感分析从研报标题拓展至摘要部分, 旨在实现从“黑箱式”综合评分向“结构化、可解释”评分的转变。相较于 ChatGPT, DeepSeek 具备成本低、性能优等优势, 使大规模文本分析更具可行性, 有助于揭示分析师情绪结构和投资信号的细节。

研报摘要内容的框架拆解: 我们系统拆解了卖方研报摘要的内容结构与分类框架, 将摘要划分为事件、点评、盈利预测与风险提示四部分, 并聚焦事件与点评部分的细分类别, 明确三大类共十小类利好因素, 包括业务发展、财务表现与其他外部因素。在此基础上, 研究引入四维度量化分析: 类别识别、情绪倾向、出现顺序与字数占比, 以更精细地提取信息结构。

模型选择与场景规则化训练: 我们采用 DeepSeek 系列模型, 为提高模型在金融场景的适用性, 通过系统化任务规则设计避免“幻想”输出, 实现信息抽取的可控性

与一致性。此外，分别实现本地部署与云端 API 调用，兼顾处理效率与硬件资源限制，为后续大规模研报分析奠定基础。通过对三种模型的测试，我们最终选择 DeepSeek-V3-671B 执行主任务，因其在输出格式规范性、结果稳定性及批量处理效率方面均优于其他版本。

结构化情绪视角下的研报摘要价值重构：分类识别方法显著提升了研报分析的颗粒度，使得我们能够从“沙里淘金”“去伪存真”“拨云见日”三个维度洞察企业真实经营价值，进而为逆势投资、剥离外部干扰和提升盈利质量判断提供有力支持。

- **情绪上的“利好密度”对未来超额收益的预测效能：**研究发现，短期内利好情绪的市场反应更为强烈，特别是“业绩超预期”和“盈利改善”类情绪信号在多个时间窗口均展现出显著的正向收益预测能力。**在个股层面**，构建“利好密度”指标，用于衡量特定时间段内，某一类别的利好情绪在全部情绪中的相对占比。基于该指标设计了主动选股策略，“盈利改善密度”策略在 20 日窗口期，Top30 组合年化收益 15.0%，相对中证 800 超额 15.6%，最大回撤 27.5%。
- **研报摘要中情绪信息的“出现顺序”与“篇幅占比”：**“出现顺序”与“篇幅占比”属于研报中利好与利空信息的“强调方式”，结果显示，利好信息出现越靠前，或篇幅占比越高，往往具有更强定价能力，尤其在战略、创新与政策类信号上更为明显；而基本面类利好则因信息普遍性，信号效力相对较弱。

研报摘要因子的综合构建与效果评估：最后基于大模型分类的研报摘要内容，综合情绪类别、篇幅占比、情绪密度等多个维度，构建了涵盖报告维度和个股维度的系列评分因子。通过引入集中度调整 (HHI) 与非线性篇幅权重，并结合类别重要性权重，逐步构建出四类核心因子：`score_mean_hhi`、`score_by_len`、`score_by_cat`、`score_by_LenCat`。结果表明，`score_by_cat` 与 `score_by_LenCat` 在多头年化收益上表现较优，而 `score_mean_hhi` 在 Rank IC 与 IC 胜率上占优。**进一步将四类因子等权合成综合评分因子 `score_report_llm`，其整体表现更为均衡，具备一定选股与收益预测能力，五分组收益结构严格单调，且与传统因子相关性较低。该因子自 2020 年以来相对中证 800 的年胜率达 100%，今年月胜率 100%，超额 10.4%；整个区间的最大回撤控制良好(相对分组等权的超额最大回撤均控制在 4% 以内，2025 年更仅为 1.2%)。**

风险提示：

量化模型基于历史数据，过去的回测业绩不代表未来；量化模型本身存在失效的风险。

重要声明

分析师声明

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，以勤勉的执业态度、专业审慎的研究方法，使用合法合规的信息，独立、客观地出具本报告，本报告所采用的数据和信息均来自市场公开信息，本人对这些信息的准确性或完整性不做任何保证，也不保证所包含的信息和建议不会发生任何变更。报告中的信息和意见仅供参考。本人过去不曾与、现在不与、未来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接收任何形式的补偿，分析结论不受任何第三方的授意或影响，特此声明。

免责声明

华安证券股份有限公司经中国证券监督管理委员会批准，已具备证券投资咨询业务资格。本报告中的信息均来源于合规渠道，华安证券研究所力求准确、可靠，但对这些信息的准确性及完整性均不做任何保证。在任何情况下，本报告中的信息或表述的意见均不构成对任何人的投资建议。在任何情况下，本公司、本公司员工或者关联机构不承诺投资者一定获利，不与投资者分享投资收益，也不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任。投资者务必注意，其据此做出的任何投资决策与本公司、本公司员工或者关联机构无关。华安证券及其所属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投资银行服务或其他服务。

本报告仅向特定客户传送，未经华安证券研究所书面授权，本研究报告的任何部分均不得以任何方式制作任何形式的拷贝、复印件或复制品，或再次分发给任何其他人，或以任何侵犯本公司版权的其他方式使用。如欲引用或转载本文内容，务必联络华安证券研究所并获得许可，并需注明出处为华安证券研究所，且不得对本文进行有悖原意的引用和删改。如未经本公司授权，私自转载或者转发本报告，所引起的一切后果及法律责任由私自转载或转发者承担。本公司并保留追究其法律责任的权利。

投资评级说明

以本报告发布之日起 6 个月内，证券（或行业指数）相对于同期沪深 300 指数的涨跌幅为标准，定义如下：

行业评级体系

- 增持—未来 6 个月的投资收益率领先沪深 300 指数 5%以上；
- 中性—未来 6 个月的投资收益率与沪深 300 指数的变动幅度相差-5%至 5%；
- 减持—未来 6 个月的投资收益率落后沪深 300 指数 5%以上；

公司评级体系

- 买入—未来 6-12 个月的投资收益率领先市场基准指数 15%以上；
- 增持—未来 6-12 个月的投资收益率领先市场基准指数 5%至 15%；
- 中性—未来 6-12 个月的投资收益率与市场基准指数的变动幅度相差-5%至 5%；
- 减持—未来 6-12 个月的投资收益率落后市场基准指数 5%至；
- 卖出—未来 6-12 个月的投资收益率落后市场基准指数 15%以上；
- 无评级—因无法获取必要的资料，或者公司面临无法预见结果的重大不确定性事件，或者其他原因，致使无法给出明确的投资评级。市场基准指数为沪深 300 指数。