

【财新网】浙商基金试水人工智能基金：排除数据噪音仍需人类智慧

现阶段的机器学习在定量分析上效率极高，但定性分析上存在局限，容易出现对噪音的过度拟合，而金融市场恰恰是噪音极大的领域。为避免这一局限，人工智能和人类智慧相结合是现阶段可行的办法

【财新网】（记者 张立君 实习记者 曹赟娴）横扫棋盘的人工智能，能否在投资领域一展身手？号称“全球首只人工智能ETF”的AIEQ在2017年推出时一度吸引了全球目光，但现在看来业绩难言乐观，成立以来累计收益仅为4.11%，跑输美股大盘。

据浙商基金梳理，人工智能技术运用在投资领域，至少包括三个不同维度的应用。

一是工具型应用，以kensho公司为代表，主要把网络爬虫、文本结构化解析等目标清晰的任务，交由AI完成；二是交易型应用，代表公司包括Winton、Citadel、D.E. Shaw等，主要完成交易执行层面的工作，尤其运用在对冲基金的短线高频交易策略中，同样也可以完全交由AI完成，可以降低成本，提升效率，降低管理费率。

真正困难的是第三类运用，即投资型应用，其主要任务是结合基本面、市场情绪、政策舆情等各个维度的大数据，进行组合管理和风险管理。代表机构有桥水基金、Two Sigma。值得注意的是，投资型应用和工具型和交易型最大的不同在于问题定义更加复杂，然而现有的AI技术尚且无法完全替代人类。

目前，浙商基金正在试水AI技术在投资型应用这一维度，“浙商智能行业优选基金”于8月26日正式首发募集，如果成功成立，可能是国内首只号称基于人工智能来管理投资组合的公募基金产品。

据查晓磊介绍，该基金与传统量化基金和大数据基金有较大差别：传统量化基金核心策略多数是多因子选股，相当于“拿一把相对固化的尺子量所有的上市公司”。而人工智能基金对于每个行业都单独用数据模型和策略模型去建模，尝试去找寻这个行业最合适的“尺子”。此外，与大数据基金相比，人工智能基金的核心竞争力更侧重策略构建的层面，而不仅仅停留在信息层面。

AI 投资的现实痛点：信噪比不佳导致模型过度拟合

在人工智能风口崛起之前，“大数据”基金一度在 2015 年-2016 年风靡一时，两年内扎堆成立 14 只。不过，随后逐渐降温，2017 年仅成立 3 只，2018 年仅成立 1 只，并逐渐销声匿迹。截至 2018 年末，现存 20 只大数据基金合计规模为 92.26 亿元，较 2015 年巅峰时期的 163.28 亿元接近腰斩。

这与业绩不佳有直接关系。在 2018 年市场单边下跌的过程中，大数据概念基金全年加权平均跌幅 25.42%，跑输上证综指和沪深 300。

大数据概念为何马失前蹄？“经过反复思考，我们认为大数据基金普遍存在几个问题，就是重数据、轻策略、无监督。”浙商基金总经理聂挺进称，这导致了大数据基金在历史回测和拟合过程中结果都很好，但是一旦实盘操作，收益和预期反而相差甚远，即出现了模型过度拟合的现象。

所谓模型过度拟合，指的是模型对历史样本数据的学习达到了很高的准确度，但是由于金融数据极低的信噪比特性，这样训练出来的模型对于样本外的数据的自适应能力非常不稳定，导致实盘业绩不佳。之所以出现过度拟合，是因为机器学习模型对外界投资信息全盘接纳，然而这个过程中模型大部分拟合的都是噪音信息。

可以说，现阶段的机器学习在定量分析上效率极高，但定性分析上存在局限。“投资这件事情的开放性决定了其问题复杂度极高，从地球到火星的距离不过是 10 的 9 次方，围棋走法的数量级是 10 的 170 次方，而仅仅 1000 个因子随机组合所构成的问题复杂度的数量级，轻轻松松便达到了 10 的 300 次方。”聂挺进感叹。

“AI+HI”：人工智能和人类的互补

可以达成共识的是，股票市场上的博弈不同于棋盘，不存在完全的信息透明，也不存在必胜策略，在多维的、非线性领域，AI 尚且无法一步到位地完全替代人类。

因此，浙商基金决定采取智能人工增强（AI+HI）的方式，即将机器算力和人类智慧相结合，以基金经理和分析师的智慧作为机器学习的对象，尽量减少信息输入阶段的噪音，从而提升学习效果。

“只要采用了机器学习，模型发生过度拟合的情况是很难避免的。”智能行

业优选的拟任基金经理向伟指出，但是也不用过度悲观，借助 AI+HI 的机制可以从两个方面来进行补充。第一是对于线上模型效果进行全面实时监控，及早发现过拟合现象，尽量控制线上模型发生过度拟合持续的时间，同时反思策略有效性，发掘新的 Alpha 来不断增强线上模型对于外界环境自适应能力；二是为了防范过度拟合现象大面积同时发生，从策略多样性角度尽可能地进行分散投资。

人工智能和人类智慧各自应该如何分工？

“对于人工智能直接做决策，需要树立谨慎的看法。但 AI 在对细分数据进行增强判断，是目前的主要方向。”聂挺进指出，规则化的、问题定义明确、训练学习结果稳定的策略将由 AI 模型独立完成，可以降低管理成本，例如运用到指数型、指数增强型产品。进一步的，利用 AI 的学习能力，可以为人的决策提供有效辅助，提高人的决策效率，也就是 AI+HI。

具体到 AI+HI 中，人工智能主要负责大数据处理、均值回归类问题，以及具有明确规律的问题。包括流动性监控、行业相对吸引力监控、估值监控，以及行业模型、交易模型、情绪模型等，监控投资线索，寻找投资标的，并对风险进行定量。

与此同时，人类基金经理则负责关注非公开数据、非均值回归的分析。核心工作是结合经验判断，辨别高品质公司和管理层，持续跟踪优质公司，总结行业规律，并对风险进行定性。

但是，聂挺进同时指出，至于“奇点突破型”领域，也就是尚未被定义的，完全没有历史可以进行学习的投资领域，仍须由人类基金经理管理。

责任编辑：蒋飞 版面编辑：刘潇