

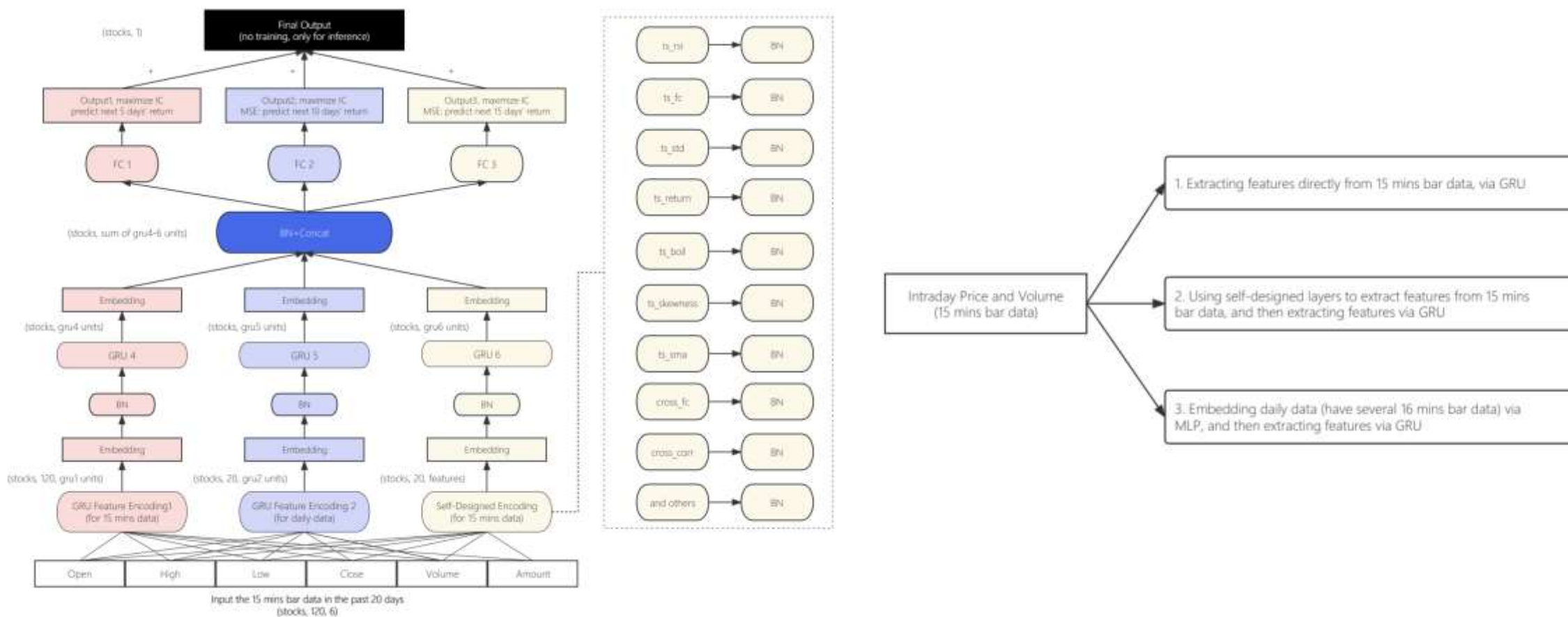
2023年 11月

诗云投资咨询

中证500指增策略介绍

基于深度学习模型的价量因子

研究动机: 1. 现实世界中存在大量非理性投资者, 其非理性行为会产生错误定价, 而该错误定价会反映在价量数据中。2. 在高频数据的交易中, 算法交易比重高+数据点多+高频数据的处理难度比低频数据高, 这使得深度学习算法在这类型数据中具有优势, 其过拟合风险相对不那么高, 数据拟合能力得以发挥。

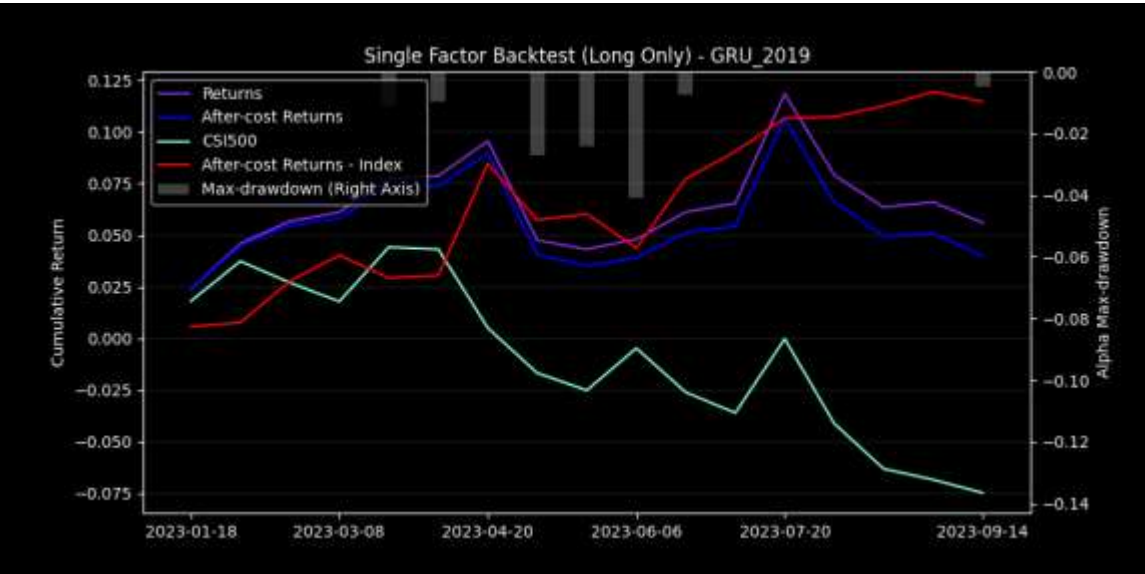


深度学习价量因子回测结果

重要批注（下同）：

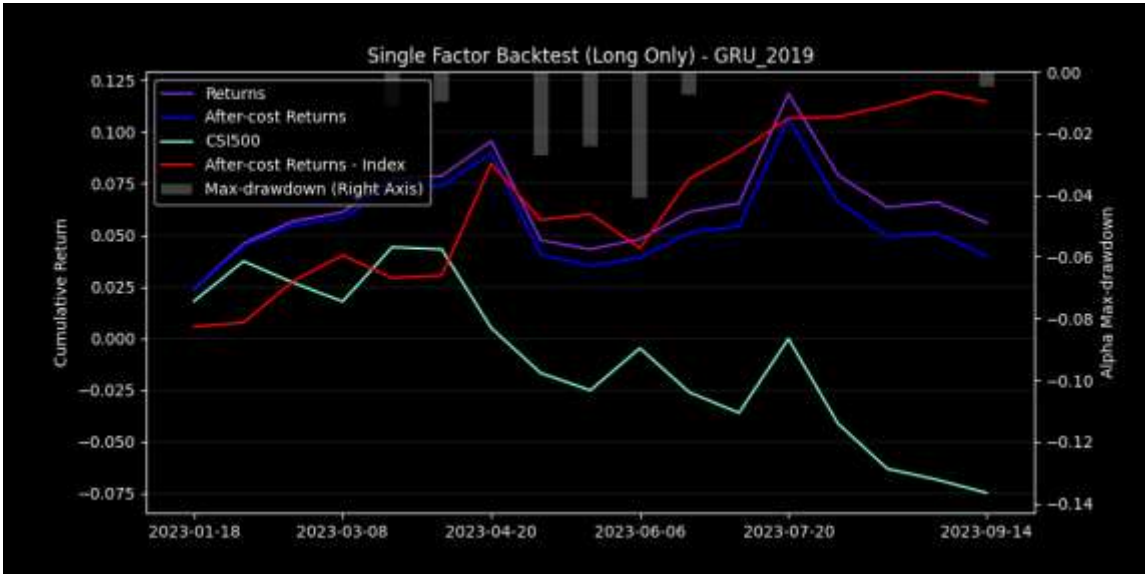
- 1. 所有交易成本设定为，买20bp，卖20bp；
- 2. 换手率为买卖的总换手率，如调仓时卖出15%再买入15%的股票，则换手率为30%；
- 3. 历史因子值已上传网盘，无需任何注册即可下载，也无限速！每个因子值的文件名为第T日，于该日收盘后计算因子值，用于T+1日的交易，交易买卖价为当日的高开低收均价；
- 4. 超额收益为（多头持仓-基准指数），并不意味着对冲基准指数，故不计算对冲成本；
- 5. 所有超额收益曲线都为累加收益曲线，并非累乘；

我们以中国大陆地区专业的机构投资者的视角来设定交易环境。历史因子值可于此处<https://cloud.189.cn/web/share?code=qe1BzyEZvaa2>（访问码：rj0v），包含必要的时间戳说明，下载无限速，免注册。



组合优化细节：

- 1. 最大化预期收益率
- 2. 中证500成分股内的持仓权重大于等于60%
- 3. 换手率约束（周频双边小于等于50%）
- 4. 严格行业中性（相对于中证500）
- 5. 相对于中证500的个股偏离小于等于1%
- 6. 沪深300+中证500内选股，一直满仓交易



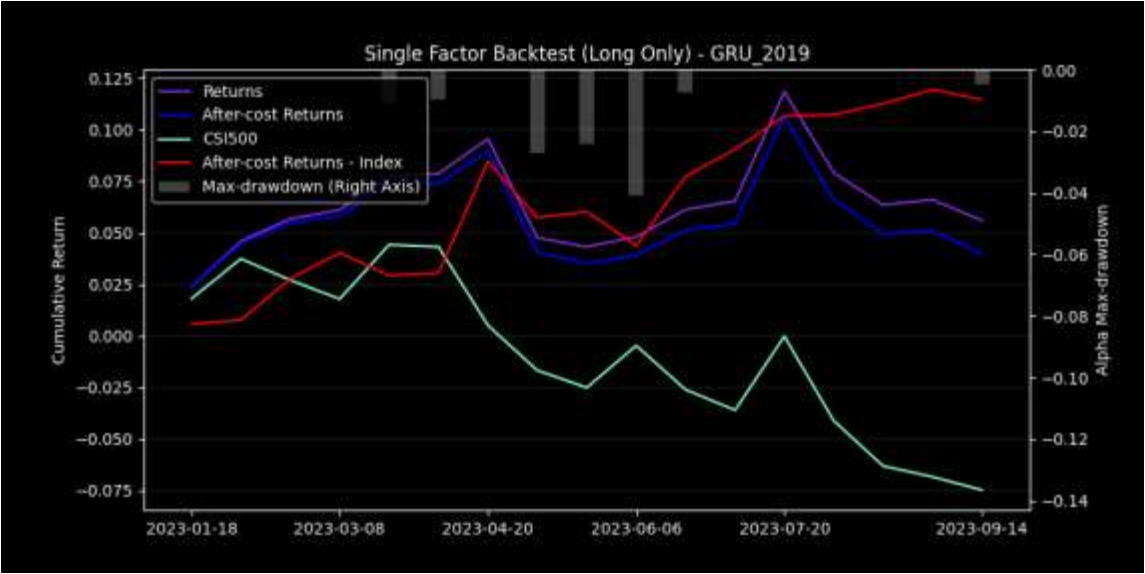
组合优化细节：

- 1. 最大化预期收益率
- 2. 无成分股持仓权重约束
- 3. 换手率约束（周频双边小于等于50%）
- 4. 严格行业中性（相对于中证500）
- 5. 相对于中证500的个股偏离小于等于1%
- 6. 沪深300+中证500内选股，一直满仓交易

业绩比较

我们基于公开数据，展示华夏中证500指数增强的业绩（该产品在公募基金同类型产品中长期居于top2），以及今年以来（20230101-20231120）公募与私募在500增强上的平均水平，并将之与我们的价量因子的表现进行对比。

数据来源：东方财富

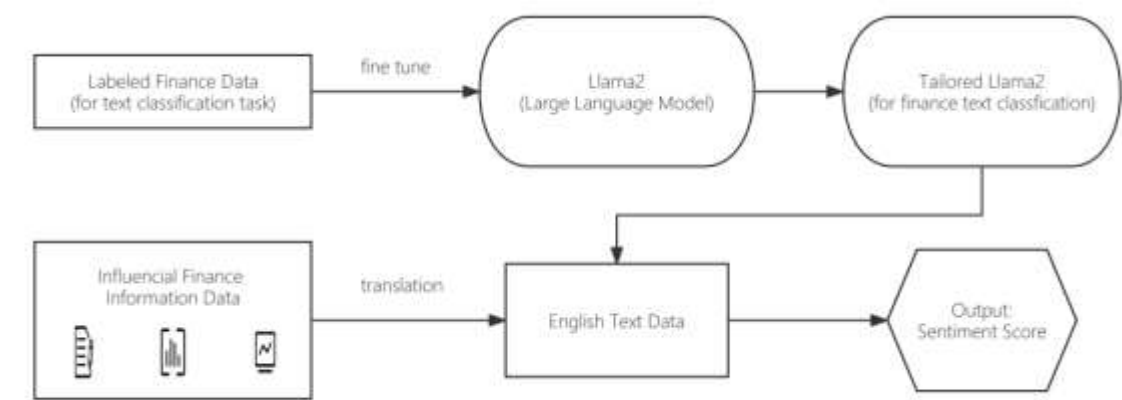


华夏中证500指数增强，其表现一直较为优异，早在2017年就开始与MSRA一起探索人工智能算法在金融投研中的应用。今年以来录得超额收益7.5%。我们单个价量因子就可以达到12.25%的超额收益，且该交易环境和换手率按照公募基金进行假设，若放在合规要求相对宽松的私募基金，则超额收益会更加可观。

基于Llama2的文本因子

研究动机: 有影响力的市场消息->市场情绪->股票的供需关系->股票价格, 我们认为<各大证券公司的研究报告>和股票市场中<高订阅量的自媒体的文章>是<有影响力的市场消息>。

我们并不依赖于或受制于任何API接口, 而是通过目前最强大的开源自然语言大模型Llama2 (对标GPT-4),在这个大语言模型的基础上, 我们用海量的自己收集的金融文本数据进行训练, 使之成为专精于金融领域的语言模型, 然后分析大量的股票文本数据, 从而构建出情绪因子。



逻辑结构:

- 1.海量金融数据+ 结构完全透明的Llama2模型->擅长当前细分领域的大语言模型
- 2.爬取各大影响力的金融信息数据->高排他性, 不拥挤的交易收益来源
- 3.中英文金融信息数据->翻译为英文文本->大语言模型进行情感打分

有影响力的金融信息:

- 1.自建网络爬虫, 爬取各大公开信息源上的有影响力的金融信息, 包括但不限于格隆汇, 中国证券报, 中国基金报, 萝卜投研, 巨潮咨询, 深交所, 上交所, 港交所等。
- 2.内部设立专业的数据清洗小组, 严格遵守算法协助清洗与人工审查追责制。



文本因子回测结果

- 重要批注（下同）：
1. 所有交易成本设定为，买20bp，卖20bp；

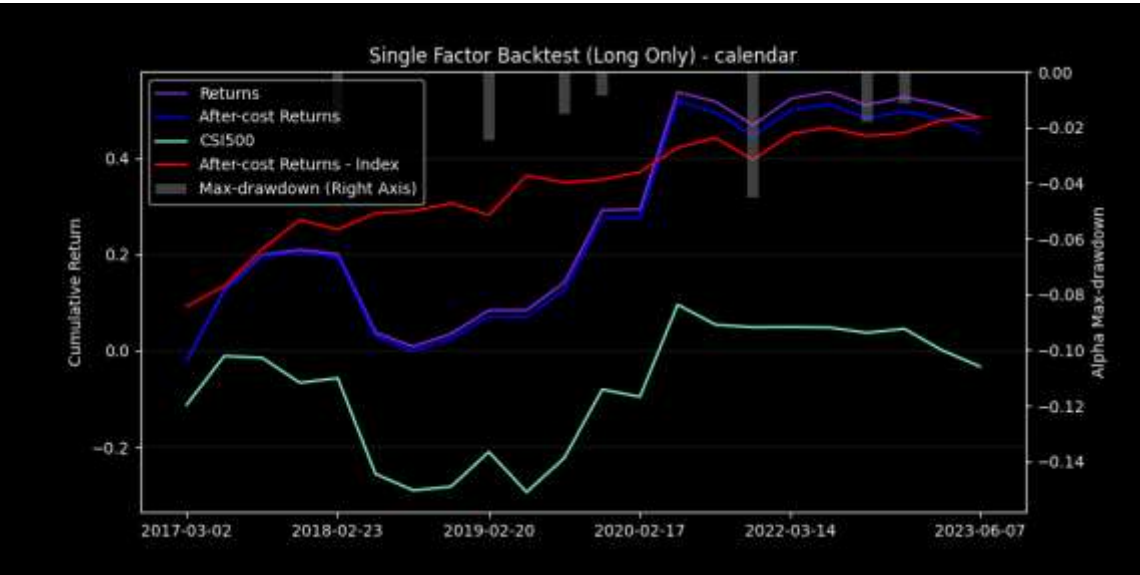
2. 换手率为买卖的总换手率，如调仓时卖出15%再买入15%的股票，则换手率为30%；

3. 历史因子值已上传网盘，无需任何注册即可下载，也无限速！每个因子值的文件名为第T日，于该日收盘后计算因子值，用于T+1日的交易，交易买卖价为当日的高开低收均价；

4. 超额收益为（多头持仓-基准指数），并不意味着对冲基准指数，故不计算对冲成本；

5. 所有超额收益曲线都为累加收益曲线，并非累乘；

我们以中国大陆地区专业的机构投资者的视角来设定交易环境。历史因子值可于此处<https://cloud.189.cn/web/share?code=blfE3erQfYzq>（访问码：ox5a）下载，包含必要的时间戳说明，下载无限速，免注册。



组合优化细节：

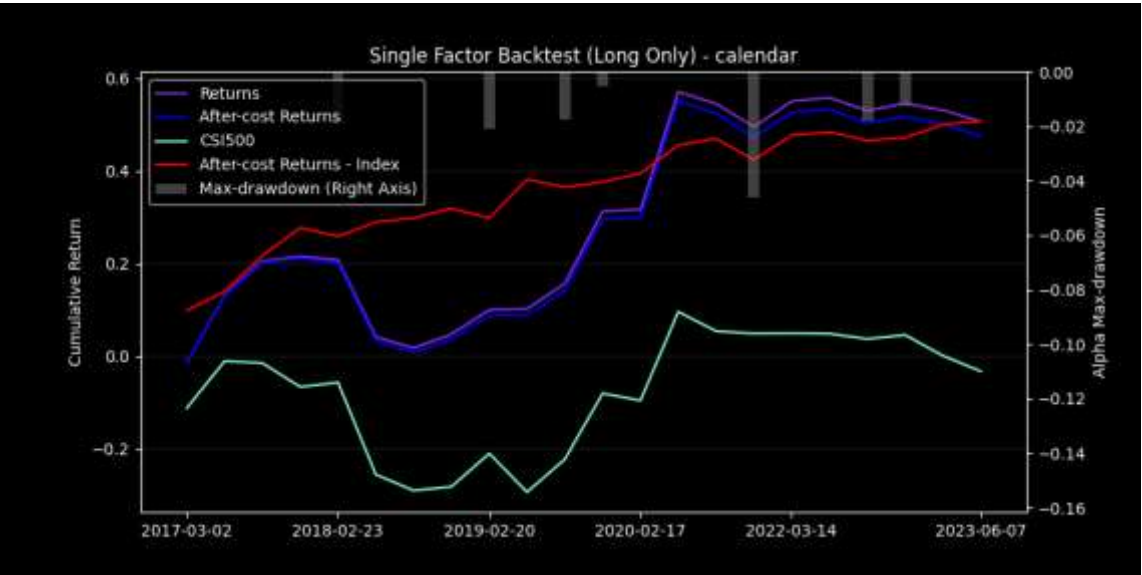
1. 最大化预期收益率

2. 中证500成分股内的持仓权重大于等于60%

3. 换手率约束（周频双边小于等于50%）
4. 严格行业中性（相对于中证500）

5. 相对于中证500的个股偏离小于等于1%

6. 沪深300+中证500内选股，一直满仓交易



组合优化细节：

1. 最大化预期收益率

2. 无成分股持仓权重约束

3. 换手率约束（周频双边小于等于50%）
4. 严格行业中性（相对于中证500）

5. 相对于中证500的个股偏离小于等于1%

6. 沪深300+中证500内选股，一直满仓交易

探索大语言模型的广阔世界



探索大语言模型的广阔世界

‘理解文本’代替‘文本分类’

文本分类属于一种半结构化数据，很多时候需要确定这条信息描述的积极点与消极点是否和我们的研究主体相关。大语言模型可以充分地去理解文本，更好地提取出结构化的观点与特征。

‘总结规律’代替‘拟合规律’

大语言模型可与symbolic regression相结合，不再是简单的拟合数据，而是将拟合的数据转化为人们可以理解的客观规律。

‘不断进化’代替‘简单统计’

参考AI Agents的设计思想，通过设置合理的规则，让大语言模型之间相互协作，监督与制衡，从而去学会如何学习，并传承记忆。学会如何学习，而不是困锁在统计指标之中，统计指标的由来与未来，皆不是由统计指标本身得来的。

平安银行公司新闻		日期/名称/语言	分类
平安银行：平安银行是一家独立的上市公司，不存在过度依赖控		2023-11-17 11:21:21	
平安银行：您的问题请咨询相应的上市公司		2023-11-16 16:51:22	
机构最新调研路线出炉：继电器龙头最受关注		2023-11-12 12:35:00	
平安银行：您的问题请咨询相应的上市公司		2023-11-09 15:12:20	
平安银行：本行未有应披露而未披露的信息		2023-11-06 15:52:12	
三季度业绩分化，北京银行还能保住“城商行一哥”宝座吗？		2023-11-04 20:07:48	
穿越市场波动，构建“稳健价值增长”，综合金融应该怎么做？		2023-11-01 05:21:00	
费、利双降的平安银行，从“站着挣钱”到“躺着挣钱”？		2023-10-31 21:17:54	
平安银行：截至2023年9月末，本行普通股股东总户数530,229户		2023-10-30 11:38:05	
平安银行：本行分红方案会综合考虑股东投资回报、监管机构对		2023-10-23 16:19:14	

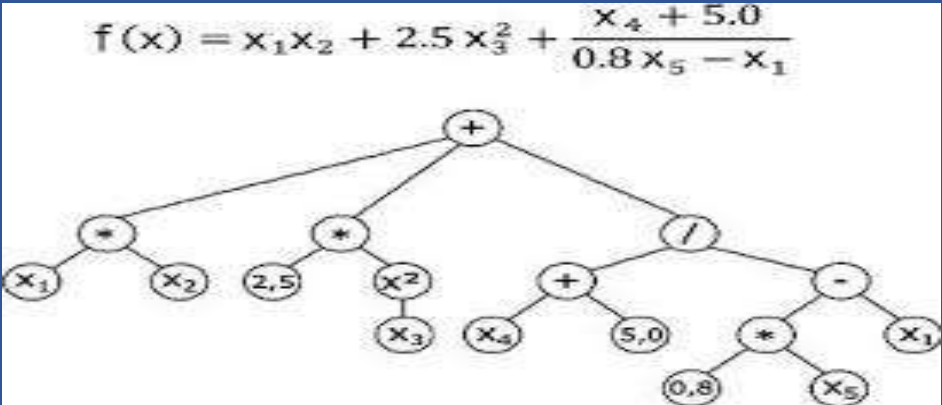


Figure 1: Generative agents create believable simulacra of human behavior for interactive applications. In this work, we demonstrate generative agents by populating a sandbox environment, reminiscent of The Sims, with twenty-five agents. Users can observe and intervene as agents they plan their days, share news, form relationships, and coordinate group activities.

定制化的研究服务

期待您与我们一起探索潜在的合作机会！

务实的态度

我们主张从根本上解决问题：1.包括弄清楚数据源的源头，对源头数据进行爬取。2.构建属于我们自己的模型，当gpt-3.5刚出来的时候，我们做不出gpt3.5，那我们就硬啃Alpaca，后面再迭代到Llama1，现在是到了Llama2（可以对标gpt-4）。我们并不会过度依赖于其他数据供应商和大语言模型接口。

不受制于是存活的根本需求，也是技术领先的基础要求。

项目可行性

我们有非常丰富的算法部署和项目经验，根据以往的经验，为客户提供算法部署或者基础设施开发服务可以极大提高客户的满意度。如您有需要，我们可以以远程控制/线下派遣员工的方式，为您提供定制化的算法部署、基础设施开发服务。如部署深度学习开发环境、开源的大语言模型微调环境等。

我们要确保这个项目在您那边的可行性， make things possible.

另类数据

探索浩如烟海的数据，去寻找可能为您带来价值的数据集。同时，我们也欢迎您给我们一些值得探索的标的，我们可以尽快的试错，就相关数据集是否有价值进行反馈。

其他市场

我们可以与您一起探索在其他市场上的投资机会，如现在估值较低、杠杆工具健全的港股；现在估值虽高，但过去长期强势，且杠杆工具健全的美股。

融通世界

这个世界是广阔且互相影响的，我们也有兴趣去获取这个世界上非常有影响力的数据，如美联储是否加息，欢迎您和我们一起去探索！

诗云投资咨询中心

利用前沿的人工智能算法、在国内外顶级买方卖方（前台核心岗位）的工作经验为合格投资者提供量化投资策略与信号。所有这些都是为了帮助我们的客户和投资组合经理驾驭金融市场，从而录得长期稳定的超额收益。

一般披露：

本材料仅供参考，并不构成向任何司法管辖区的任何人购买或出售任何证券的投资建议、推荐或要约或招揽。

该信息并不完整或详尽，并且对于此处所含信息的准确性或完整性不作任何明示或暗示的陈述或保证。所表达的意见截至2023年11月，如有更改，恕不另行通知。读者可自行决定是否信赖本材料中的信息。投资涉及风险。

本材料可能包含估计和前瞻性陈述，其中可能包括预测，但并不代表对未来业绩的保证。历史并不能完全重演，模型存在失效的风险。所有投资都涉及风险，包括可能损失本金。

截止2023年11月，该材料仅在中国大陆和中国香港地区发行，并遵守发行地的证券法，如与任何正在发生改变的规定相违背，我们将立刻做出整改，并保留对该材料连带的责任与风险的最终解释权。潜在投资者应自行了解其国籍国、居住国或住所国可能相关的任何适用的法律要求以及税收和外汇管制法规。

©2023 诗云投资咨询中心或其附属公司版权所有。

联系方式：

方颖，Tel: (+86) 15902054135, (+852) 5105 3158, E-mail: fangjie@poem-cloud.com.cn