

行业点评

无损长文本长度达200万字，Kimi打开大模型应用新世界

强于大市（维持）

行情走势图



相关研究报告

【平安证券】行业点评*计算机*Anthropic推出Claude 3，能力对标GPT-4*强于大市20240305

【平安证券】行业点评*计算机*政府工作报告提出人工智能+，持续关注AI投资机会*强于大市20240305

【平安证券】行业点评*计算机*OpenAI发布文生视频大模型Sora，在全球视频大模型领域取得里程碑式进展*强于大市20240218

证券分析师

付强 投资咨询资格编号
S1060520070001
FUQIANG021@pingan.com.cn

闫磊 投资咨询资格编号
S1060517070006
YANLEI511@pingan.com.cn

黄韦涵 投资咨询资格编号
S1060523070003
HUANGWEIHAN235@pingan.com.cn



事项：

近日，国内人工智能公司月之暗面（Moonshot AI）在其微信公众号“Moonshot AI”宣布，该公司的Kimi智能助手在长上下文窗口技术上取得突破，无损上下文长度可以到200万字。此前，在2023年10月公司的智能助手可以实现20万无损上下文长度，最新能力提升了一个数量级。

平安观点：

- **Kimi智能助手短期内能力得到了大幅提升，在国内热度快速提高。**此次发布的无损上下文长度达到200万字，较此前的20万字，中间仅相隔了5个月左右。而公司月之暗面也是在2023年4月份才成立，成立时间也不足一年。Kimi智能助手又称Kimi Chat，是月之暗面基于自研千亿参数大模型打造的对话式AI助手产品，于2023年11月正式面向市场应用。该产品最强的能力在于长上下文处理，包括长文总结和生成、联网搜索、数据处理、编写代码、用户交互和翻译等。上线之后，该工具的热度快速提升。根据Similarweb数据显示，近几周Kimi的访问量上升趋势非常明显，该网站统计的最近四周（2.20-2.26、2.27-3.4、3.5-3.11、3.12-3.18）的访问量分别是100.3万、112.8万、152.0万和225.0万。虽然公司对服务器持续进行了扩容，但是面对快速的用户增长，压力已经开始显现。
- **长文本有望打开大模型应用新世界。**大模型的参数量决定了其能够支持多复杂的“计算”，而能够接收多少文本输入（即长文本技术）则决定了大模型有多大的“内存”，两者共同决定模型的应用效果。当前大模型输入长度普遍较低的现状对其技术落地产生了极大制约，比如虚拟角色会“忘记”部分重要信息，Agent无法获得全面的输入信息可能出现运行失败，部分游戏产品由于无法处理长文本能力而被迫简化剧情等。Kimi支持更长的上下文意味着大模型拥有更大的“内存”，从而使得大模型的应用更加深入和广泛，公司官方公众号显示，Kimi可以通过多篇财报进行市场分析、处理超长的法务合同、快速梳理多篇文章或多个网页的关键信息、基于长篇小说设定进行角色扮演等等。大模型无损上下文长度的数量级提升，未来会帮助用户打开对AI应用场景的想象力，包括完整代码库的分析理解、可以自主完成多步骤复杂任务的智能体Agent、不会遗忘关键信息的终身助理、真正统一架构的多模态模型等。
- **“无损压缩”和文本长度提升是长文本技术需要兼顾的要求。**公司创始人表示，如果要实现通用人工智能，无损的长上下文将会是一个很关键的基础技术。历史上所有的模型架构演进，本质上都是在提升有效的、无损的上下文长度。上下文长度在提升的过程中，需要兼顾长度和无损压缩水平两个指标，这样才能有意义的规模化。从此次升级的间隔来看，

时间非常短，可见公司并没有走渐进式的迭代路线，当然面临的技术难度应该也更大。公司的研发和技术团队从模型预训练到对齐、推理环节均进行了原生的重新设计和开发，在千亿参数下实现了无损的长程注意力机制，不依赖于滑动窗口、降采样、小模型等对性能损害较大的“捷径”方案，兼顾了对长度和“无损”的两项指标。

投资建议：Kimi的火热，是国内在大模型基础技术研究方面的突破的一个缩影。无损长文本模型的出现，也解决了很多大模型在应用中的痛点，打开了大模型的应用空间。AIGC最终的落地，还是要找到合适的场景。Kimi作为千亿大模型，能够支持复杂运算，同时还能接受和处理大文本，解决了很多大模型实际应用中的难题，后续商业化潜力将凸显。目前，Kimi智能助手已在苹果iOS应用、Android应用、小程序、网页Web等多端平台上线。我们继续看好AIGC产业链，尤其是大模型的应用潜力。在标的方面：1) 算力方面，推荐浪潮信息、中科曙光、紫光股份、神州数码、海光信息、龙芯中科，建议关注工业富联、寒武纪、景嘉微、高新发展、拓维信息、软通动力；2) 算法方面，推荐科大讯飞；3) 应用场景方面，强烈推荐中科创达、恒生电子、盛视科技，推荐金山办公，建议关注万兴科技、同花顺、彩讯股份；4) 网络安全方面，强烈推荐启明星辰。

风险提示：1) 国内算力提升速度不及预期的风险；2) 版权等合规监管风险；3) 技术演进风险。

平安证券研究所投资评级:

股票投资评级:

强烈推荐（预计6个月内，股价表现强于市场表现20%以上）
推荐（预计6个月内，股价表现强于市场表现10%至20%之间）
中性（预计6个月内，股价表现相对市场表现在±10%之间）
回避（预计6个月内，股价表现弱于市场表现10%以上）

行业投资评级:

强于大市（预计6个月内，行业指数表现强于市场表现5%以上）
中性（预计6个月内，行业指数表现相对市场表现在±5%之间）
弱于大市（预计6个月内，行业指数表现弱于市场表现5%以上）

公司声明及风险提示:

负责撰写此报告的分析师(一人或多人)就本研究报告确认:本人具有中国证券业协会授予的证券投资咨询执业资格。

平安证券股份有限公司具备证券投资咨询业务资格。本公司研究报告是针对与公司签署服务协议的签约客户的专属研究产品，为该类客户进行投资决策时提供辅助和参考，双方对权利与义务均有严格约定。本公司研究报告仅提供给上述特定客户，并不面向公众发布。未经书面授权刊载或者转发的，本公司将采取维权措施追究其侵权责任。

证券市场是一个风险无时不在的市场。您在进行证券交易时存在赢利的可能，也存在亏损的风险。请您务必对此有清醒的认识，认真考虑是否进行证券交易。

市场有风险，投资需谨慎。

免责条款:

此报告旨在发给平安证券股份有限公司(以下简称“平安证券”)的特定客户及其他专业人士。未经平安证券事先书面明文批准，不得更改或以任何方式传送、复印或派发此报告的材料、内容及其复印本予任何其他人。

此报告所载资料的来源及观点的出处皆被平安证券认为可靠，但平安证券不能担保其准确性或完整性，报告中的信息或所表达观点不构成所述证券买卖的出价或询价，报告内容仅供参考。平安证券不对因使用此报告的材料而引致的损失而负上任何责任，除非法律法规有明确规定。客户并不能仅依靠此报告而取代行使独立判断。

平安证券可发出其它与本报告所载资料不一致及有不同结论的报告。本报告及该等报告反映编写分析员的不同设想、见解及分析方法。报告所载资料、意见及推测仅反映分析员于发出此报告日期当日的判断，可随时更改。此报告所指的证券价格、价值及收入可跌可升。为免生疑问，此报告所载观点并不代表平安证券的立场。

平安证券在法律许可的情况下可能参与此报告所提及的发行商的投资银行业务或投资其发行的证券。

平安证券股份有限公司2024版权所有。保留一切权利。

平安证券

平安证券研究所

电话:4008866338

深圳

深圳市福田区益田路5023号平安金融中心B座25层

上海

上海市陆家嘴环路1333号平安金融大厦26楼

北京

北京市丰台区金泽西路4号院1号楼丽泽平安金融中心B座25层