

30 May 2023

Popular science report of AI: Algorithms, Computing Power, Data and Applications Outperform (Maintain)

Research Analyst
Ruibin Chen

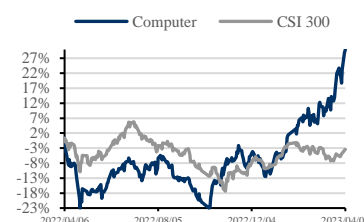
(852) 3982 3212

chenrobin@dwzq.com.hk

Investment Thesis

- **As a new paradigm of AI development, big models are a milestone for artificial intelligence towards general intelligence.** Big models refer to models that can adapt to a series of downstream tasks after training on large-scale general data, which are essentially based on statistical language models, but the "emergent capability" gives them powerful reasoning capabilities. The existing big model framework is essentially the same, and almost all large language models with parameter scales exceeding 100 billion adopt the GPT mode. However, different types of enterprises give their own advantages in the field they are in, and the developed big models still have differences in function. Technology plays a decisive role in the effect of big models, so the future competition pattern also depends on technological breakthroughs.
- **Calculating power is the "oil" of the AI era.** The training and inference of big models will use the calculating power support of AI chips. Under the same data and algorithms, calculating power is the key to the development of big models, and it is the "oil" of the AI era. Assuming one-month training time of GPT-3, we estimate that 843 Nvidia A100 chips are needed. We calculate that ~16255 Nvidia A100 chips are needed if the DAU of GPT-3 reach 50 million. As GPT-4 uses multimodal data, we estimate that its calculating power demand is >10 times of GPT-3. China's leading internet companies have successively invested in big models, and we expect that only ten leading factories may add ~200,000 A100 demand in one year. In the long run, the demand is expected to exceed 2 million chips, and the new calculating power demand is likely to double the size of the calculating power market. In 2021, Nvidia occupied more than 80% of the market share in the China acceleration card market, and the performance of domestic AI chips is still lower than that of foreign ones. We expect the launch of domestic big models to drive the development of domestic chips.
- **Data resources are one of the important driving forces for the development of the AI industry.** Datasets, as the core component of data resources, refer to data specially designed, collected, cleaned, labeled, and managed for the training of AI algorithm models. The performance of large-scale language models is strongly dependent on parameter scale N, dataset size D and computing power C. The training data mainly comes from Wikipedia, books, journals, Reddit social news sites, Common Crawl, and other datasets. GPT4 relies on a large amount of multimodal data for training. In the future, the competitiveness of AI models may be reflected in the quality and scarcity of data. The development of data elements market can promote the further opening of relevant public, corporate and individual data, providing important support for China's AI development.
- **AI empowers all walks of life, and the future is promising for AI applications.** AI is comparable to the fourth technological revolution, and the most direct application for this time is in the field of content creation, opening the imagination boundary of the industry. We should look for areas where the application functions are significantly improved, customer stickiness is significantly improved, and the market size is greatly increased with the empowerment of AI, mainly including content creation, office software, ERP, robots, and chip design. At present, some big model manufacturers have started industrial applications, but computing power is still the main reason for limiting large-scale commercialization of AI. Once solved, the information industry may be directly benefited from AI+, so we are optimistic for leaders in the information industry.
- **Recommendations:** Algorithms: we recommend large-model companies that have a first-mover advantage, such as 360 Security Tech, Iflytek, Hithink Royalflush. Computing power: we recommend Changsha Jingjia Micro, Changsha Jingjia Micro, Digital China. Data: we recommend leading information companies in each sub-track, such as Sichuan Jiuyuan Yinhai Software, Anhui Ronds Science&Tech, and Zhejiang Supcon Tech. Applications: we recommend companies with "killer" application potential, such as Beijing Kingsoft Office Software, Yonyou Network Tech, and Hundsun Tech.
- **Risks:** Lower-than-expected policy implementation, Intensified industry competition.

Industry performance



Related reports

《华为盘古大模型产业链梳理》

2023-03-27

《数据安全，为数据要素市场发展保驾护航》

2023-03-24

Soochow Securities International Brokerage Limited would like to acknowledge the contribution and support provided by Soochow Research Institute, and in particular its employees **Zijing Wang (王紫敬)** and **Shijie Wang (王世杰)**.

计算机行业深度报告

AI 偏向科普性报告：围绕算法、算力、数据和应用 增持（维持）

2023 年 5 月 30 日

分析师
陈睿彬

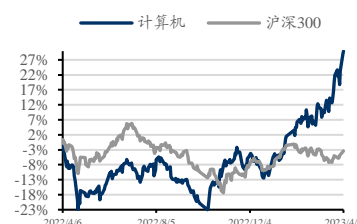
(852) 3982 3212

chenrobin@dwzq.com.hk

投资要点

- **大模型是 AI 开发的新范式，是人工智能迈向通用智能的里程碑：**大模型指通过在大规模宽泛的数据上进行训练后能适应一系列下游任务的模型，本质依旧是基于统计学的语言模型，只不过“突现能力”赋予其强大的推理能力。现有的大模型的框架在本质上是统一的，几乎所有参数规模超过千亿的大语言模型都采取 GPT 模式，但是不同类型的企业给予自己所在领域的优势，开发的大模型在功能上还是有所差异。技术对大模型的效果具有决定作用，因此未来竞争格局也依赖于技术突破。
- **算力是 AI 时代的“石油”：**大模型的训练和推理都会用到 AI 芯片的算力支持，在数据和算法相同情况下，算力是大模型发展的关键，是人工智能时代的“石油”。我们假设 GPT-3 训练时间为一个月，则需要 843 颗英伟达 A100 芯片。我们假设 GPT-3 每日日活为 5000 万，则需要约 16255 颗英伟达 A100 芯片。GPT-4 为多模态数据，我们预计算力需求是 GPT-3 的 10 倍以上。中国大厂相继布局大模型，我们测算，仅十家头部厂商大模型 1 年内有望增加约 20 万片 A100 需求量。长期来看，则需求量有望超 200 万片，新增算力需求将使算力市场增长 2 倍以上。2021 年，中国加速卡市场中 Nvidia 占据超过 80% 市场份额，国产 AI 芯片性能与海外仍有差距，国产大模型推出有望加快国产芯片发展。
- **数据资源是 AI 产业发展的重要驱动力之一：**数据集作为数据资源的核心组成部分，是指经过专业化设计、采集、清洗、标注和管理，生产出来的专供人工智能算法模型训练的数据。大规模语言模型性能强烈依赖于参数规模 N，数据集大小 D 和计算量 C，训练数据主要来自于维基百科、书籍、期刊、Reddit 社交新闻站点、Common Crawl 和其他数据集，GPT4 依靠大量多模态数据训练。未来 AI 模型的竞争力或体现在数据质量和稀缺性，发展数据要素市场，促进相关公共、企业、个人数据的进一步放开，将为国内 AI 发展提供重要支撑。
- **AI 赋能各行各业，未来是 AI 应用的星辰大海：**AI 堪比第四次技术革命，本轮最直接的应用在内容创作领域，打开产业的想象边界。我们应该去寻找在 AI 赋能下，应用功能显著改善、客户粘性显著提升，市场空间大幅提升的领域，主要有内容创作，办公软件，ERP，机器人以及芯片设计领域。当前部分大模型厂商已经开启产业化应用，但是算力依旧是限制 AI 大规模商业化落地的主要原因，一旦解决，直接受益 AI+ 的将是信息化行业，因此我们看好各行业信息化领域处于优势地位的龙头公司。
- **投资建议：算法上，我们建议关注已经有先发优势的大模型公司：**三六零、科大讯飞、同花顺等，此外还有一些实施企业，如软通动力、润和软件、汉得信息等；**算力上，我们推荐景嘉微、中科曙光、神州数码，**建议关注海光信息、寒武纪、四川长虹、拓维信息等；**数据上，我们推荐各细分赛道的信息化龙头企业，**如久远银海、容知日新、中控技术，建议关注国能日新、千方科技等；**应用上，我们推荐在具备“杀手级”应用潜能的厂商**金山办公、用友网络、恒生电子，建议关注广联达、石基信息等。
- **风险提示：**政策推进不及预期；行业竞争加剧

行业走势



相关研究

《华为盘古大模型产业链梳理》

2023-03-27

《数据安全，为数据要素市场发展保驾护航》

2023-03-24

特此致谢东吴证券研究所对本报告专业研究和分析的支持，尤其感谢王紫敬和王世杰的指导。

1. 算法：大模型——人工智能迈向通用智能的里程碑

当前大模型的商业模式是“通用大模型+产业模型”。底层 AI 大模型的研发具有极高的研发门槛，面临高昂的成本投入，不利于人工智能技术在千行百业的推广。而具有数据、算力、算法综合优势的企业可以将模型的复杂生产过程封装起来，通过低门槛、高效率的生产平台，向千行百业提供大模型服务。

一旦大模型技术出现突破，行业竞争格局有望一家独大。由于当前大模型的技术是公用的，没有哪家存在明显的技术领先，因此各行各业厂商纷纷入局，希望分得一杯羹，因此出现了大模型百花齐放的竞争格局。一旦出现技术突破，大模型的准确度以及智能化出现了“碾压”的优势，行业需求会迅速向 NO.1 集中，有望形成一家独大的竞争格局。

2. 算力：AI 训练的基础设施

国产 AI 芯片性能与海外仍有差距，国产大模型推出有望加快国产芯片发展。国产 AI 芯片厂商主要有寒武纪，景嘉微，沐曦，燧原等，但其产品性能距离海外仍有差距。以国产寒武纪为例，MLU370 性能为 FP32 24TFLOPS，仅为英伟达 A100 的 10% 不到。要达到同等算力要求，国产芯片片数需求量会更大，但大量 AI 芯片并行运行会对控制能力有较高要求，难以满足。但发展自己的 AI 芯片产业迫在眉睫，各家厂商正在快速追赶。

3. 数据：AI 发展的驱动力

高质量、稀缺的数据放开对 AI 发展至关重要。发展国内自己的大模型需要国内的高质量、稀缺数据。然而，根据发改委高技术司，我国政府数据资源占全国数据资源的比重超过 3/4，开放的规模却不足美国的 10%，个人和企业可以利用的规模更是不及美国的 7%，但这类数据的开放共享程度不高，全国开放数据集规模仅约为美国的 11%，数据有待进一步开放汇集，为开发更符合国内需求的大模型提供基础。发展数据要素市场，促进相关公共、企业、个人数据的进一步放开，将为国内 AI 发展提供重要支撑。

我们可以主要关注两个方面：能够采集、处理细分行业稀缺数据的厂商：久远银海、容知日新、国能日新、千方科技、中控技术、千方科技、用友网络等，以及具有专业数据处理服务能力的通用第三方厂商：海天瑞声等。

4. 应用：AI 的星辰大海

我们应该去寻找“杀手级”的下游应用市场。所谓“杀手级”应用市场，即在 AI 赋能下，该应用功能显著改善，客户粘性显著提升，最后体现为客户付费率和付费单价显著提升，市场空间大幅提升。我们认为根据美国产业发展现状来看，当前来看已经涌现的“杀手级”应用领域主要有内容创作，办公软件，ERP，机器人以及芯片设计领域。

最直接的应用在内容创作领域。ChatGPT 的功能核心是基于文本的理解和分析，与内容创作行业趋同。ChatGPT 可用于创建新闻文章、博客文章甚至小说等内容，它可以生成原创且连贯的内容，为内容创作者节省时间和资源。

在办公软件领域是划时代的生产力的解放。我们认为办公场景是当前所能看到的 AI 最大级别应用场景，Microsoft 365 Copilot 将会带来需求的刚性，是人工智能杀手级应用。在 Word 中，Copilot 可以帮助打草稿、排版、修改；在 Excel 中，Copilot 可以帮助用户修改样式，并进行数据分析、预测、可视化等；在 PPT 中，Copilot 能够基于文本自动绘制 PPT，根据 PPT 生成讲稿。

与 ERP 的结合有望重构企业管理。ERP 能够学习企业管理“通用数据”，又能学习企业管理“私有数据”，让 ERP 在做到贯彻领先企业管理理念的同时，越来越个性化，低成本满足企业对 ERP 定制化的需求。ERP 使用流程繁琐，使用 ChatGPT 和直接询问，获得想要的信息，可降低使用者门槛。

ChatGPT 解决了机器人的痛点。ChatGPT 开启了一种新的机器人范式，允许潜在的非技术型用户参与到回路之中，ChatGPT 可以为机器人场景生成代码。在没有任何微调的情况下，利用 LLM 的知识来控制不同的机器人动作，以完成各种任务。

看好 ChatGPT 在芯片设计领域的应用。传统的芯片设计强烈依赖模板而忽视了大量可以复用的优秀数据，同时数据量大导致 ChatGPT 泛化性更好。此外芯片硬件模块相对单一，有一些成熟范式，芯片设计代码复杂但人工不足，这些都与 ChatGPT 有很好的互补。

5. 投资建议与相关标的

算法上，我们建议关注已经有先发优势的大模型公司：三六零、科大讯飞、同花顺等，此外还有一些实施企业，如软通动力、润和软件、汉得信息等；

算力上，我们推荐景嘉微、中科曙光、神州数码，建议关注海光信息、寒武纪、四川长虹、拓维信息等；

数据上，我们推荐各细分赛道的信息化龙头企业，如久远银海、容知日新、中控技术，建议关注国能日新、千方科技等；

应用上，我们推荐在具备“杀手级”应用潜能的厂商金山办公、用友网络、恒生电子，建议关注广联达、石基信息等。

6. 风险提示

政策推进不及预期。相关政策推进受到多种因素影响，节奏和力度可能不及预期。

行业竞争加剧。行业市场空间广阔，可能吸引更多公司参与行业竞争。

分析师声明:

本人以勤勉、独立客观之态度出具本报告，报告所载所有观点均准确反映本人对于所述证券或发行人之个人观点；本人于本报告所载之具体建议或观点于过去、现在或将来，不论直接或间接概与本人薪酬无关。过往表现并不可视作未来表现之指标或保证，亦概不会对未来表现作出任何明示或暗示之声明或保证。

利益披露事项:

刊发投资研究之研究分析员并不直接受投资银行或销售及交易人员监督，并不直接向其报告。

研究分析员之薪酬或酬金并不与特定之投资银行工作或研究建议挂钩。

研究分析员或其联系人并未从事其研究 / 分析 / 涉及范围内的任何公司之证券或衍生产品的买卖活动。

研究分析员或其联系人并未担任其研究 / 分析 / 涉及范围内的任何公司之董事职务或其他职务。

东吴证券国际经纪有限公司（下称「东吴证券国际」或「本公司」）或其集团公司并未持有本报告所评论的发行人的市场资本值的 1% 或以上的财务权益。

东吴证券国际或其集团公司并非本报告所分析之公司证券之市场庄家。

东吴证券国际或其集团公司与报告中提到的公司在最近的 12 个月内没有任何投资银行业务关系。

东吴证券国际或其集团公司或编制该报告之分析师与上述公司没有任何利益关系。

免责声明

本报告由东吴证券国际编写，仅供东吴证券国际的客户使用，本公司不因接收人收到本报告而视其为客户。如任何司法管辖区的法例或条例禁止或限制东吴证券国际向收件人提供本报告，收件人并非本报告的目标发送对象。本研究报告的编制仅供一般刊发，并不考虑接收本报告之任何特定人士之特定投资对象、财政状况、投资目标及特别需求，亦非拟向有意投资者作出特定证券、投资产品、交易策略或其他金融工具的建议。阁下须就个别投资作出独立评估，于作出任何投资或订立任何交易前，阁下应征求独立法律、财务、会计、商业、投资和/或税务意见并在做出投资决定前使其信纳有关投资符合自己的投资目标和投资界限。本报告应受香港法律管辖并据其解释。

本报告所载资料及意见均获自或源于东吴证券国际可信之数据源，但东吴证券国际并不就其准确性或完整性作出任何形式的声明、陈述、担保及保证(不论明示或默示)，于法律及 / 或法规准许情况下，东吴证券国际概不会就本报告所载之资料引致之损失承担任何责任。本报告不应倚赖以取代独立判断。

本报告所发表之意见及预测为于本报告日的判断，并可予更改而毋须事前通知。除另有说明外，本报告所引用的任何业绩数据均代表过往表现，过往表现亦不应作为日后表现的可靠预示。在不同时期，东吴证券国际可能基于不同假设、观点及分析方法发出与本报告所载意见、评估及预测不一致的研究报告。分析中所做的预测收益可能基于相应的假设，任何假设的改变可能会对本报告预测收益产生重大影响，东吴证券国际并不承诺或保证任何预测收益一定会实现。

东吴证券国际的销售人员、交易人员以及其他专业人士可能会依据不同假设和标准、采用不同的分析方法而口头或书面表达与本报告意见及建议不一致的市场评论和/或交易观

点, 本公司没有将此意见及建议向报告所有接收者进行更新的义务。东吴证券国际及其集团公司的各业务部门, 如有投资于本报告内所涉及的任何公司之证券或衍生产品时, 其所作出的投资决策可能与本报告所述观点并不一致。

本报告及其所载内容均属机密, 仅限指定收件人阅览。本报告版权归东吴证券国际所有, 未经本公司同意, 不得以任何方式复制、分发或使用本报告中的任何资料。本报告仅作参考用途, 任何部分不得在任何司法管辖权下的地方解释为提呈或招揽购买或出售任何于报告或其他刊物内提述的任何证券、投资产品、交易策略或其他金融工具。东吴证券国际毋须承担因使用本报告所载数据而可能直接或间接引致之任何责任, 损害或损失。

特别声明

东吴证券国际可能在法律准许的情况下参与及投资本报告所述证券发行人之融资交易, 也可能为有关发行人提供投资银行业务服务或招揽业务, 及 / 或于有关发行人之证券或期权或其他有关投资中持仓或持有权益或其他重大权益或进行交易。阁下应考虑到东吴证券国际及/或其相关人员可能存在影响本报告及所载观点客观性的潜在利益冲突, 请勿将本报告视为投资或其他决策的唯一信赖依据。

本报告及其所载信息并非针对或意图发送给任何就分派、刊发、可得到或使用此报告而导致东吴证券国际违反当地法律或规例或可致使东吴证券国际受制于相关法律或规例的任何地区、国家或其他司法管辖区域的公民或居民。接收者须自行确保使用本报告符合当地的法律及规例。

评级标准:

公司投资评级:

- 买入: 预期未来 6 个月个股涨跌幅相对大盘在 15% 以上;
- 增持: 预期未来 6 个月个股涨跌幅相对大盘介于 5% 与 15% 之间;
- 中性: 预期未来 6 个月个股涨跌幅相对大盘介于 -5% 与 5% 之间;
- 减持: 预期未来 6 个月个股涨跌幅相对大盘介于 -15% 与 -5% 之间;
- 卖出: 预期未来 6 个月个股涨跌幅相对大盘在 -15% 以下。

行业投资评级:

- 增持: 预期未来 6 个月内, 行业指数相对强于大盘 5% 以上;
- 中性: 预期未来 6 个月内, 行业指数相对大盘 -5% 与 5%;
- 减持: 预期未来 6 个月内, 行业指数相对弱于大盘 5% 以上。

东吴证券国际经纪有限公司

Level 17, Three Pacific Place, 1 Queen's Road East, Hong Kong

香港皇后大道东 1 号太古广场 3 座 17 楼

Tel 电话: (852) 3983 0888(公司) (852) 3983 0808 (客户服务)

公司网址: <http://www.dwzq.com.hk/>

