

2024年09月08日



华鑫证券
CHINA FORTUNE SECURITIES

英特尔发布酷睿 Ultra 200V 系列处理器, 开源模型 Reflection 70B 发布

—计算机行业周报

推荐(维持)

投资要点

分析师: 宝幼琛 S1050521110002
baoyc@cfsc.com.cn

行业相对表现

表现	1M	3M	12M
计算机(申万)	-3.1	-10.9	-35.7
沪深300	-3.3	-9.6	-13.6

市场表现



资料来源: Wind, 华鑫证券研究

相关研究

- 1、《计算机行业周报: AMD RDNA4 GPU 架构显卡曝光, 智谱发布基座大模型 GLM-4-Plus》2024-09-01
- 2、《计算机行业周报: Akeana 推出新的 RISC-V 芯片设计, xAI 发布 Grok-2 大模型》2024-08-18
- 3、《计算机行业深度报告: 政策东风至, 车路云发展正当时》2024-08-13

算力: 算力租赁价格平稳, 英特尔发布酷睿 Ultra 200V 系列处理器

英特尔正式发布了酷睿 Ultra 200V 系列处理器, 首批搭载这款处理器的笔记本电脑在 9 月 24 日起发售, 现在已经开启预售, 包括宏碁、华硕、戴尔、惠普、联想、LG、微星和三星在内的 80 款 PC 产品。

该处理器为有史以来最高效的 x86 处理器, 对比前代 Meteor Lake, 整体功耗降低达 50%; 实现 x86 平台最佳电池续航, 在生产力用例中续航长达 20 小时; 每线程处理性能比前代 Meteor Lake 提高 3 倍, 峰值性能提高 8 倍。

AI 算力大幅提升, CPU+GPU+NPU 算力突破 120TOPS, 是 Meteor Lake 的 10 倍; NPU 数量从两颗增加到四颗, 单独的 NPU 4 算力达 48TOPS, 是 Meteor Lake 的 4 倍; 首次亮相的 X(e)2 图形微架构, 在移动图形性能方面带来平均 30% 的性能提升

Ultra 200V 系列处理器采用全新的 P-Core Lion Cove 架构和 E-Core Skymont 架构, 性能核与能效核的调度逻辑进行了改变, 以更好地分配性能与功耗, 同时集成了独显架构看齐的 Xe2 核显; 计算模块采用台积电 N3B 工艺代工, 平台控制模块采用 N6 工艺, 首次将 LPDDR5x 内存与计算模块封装在一起, 类似智能手机的 SoC 结构。

英特尔表示, 在相同性能表现下, Ultra 9 288V 比高通 X Elite 的旗舰处理器 X1E-80-100 功耗低 40%, 与苹果 M3 芯片旗鼓相当。

AI 应用: 开源模型新黑马 Reflection 70B 发布, 使用自我纠错技术

当地时间 9 月 6 日, Reflection 70B 模型得到发布, 模型表现达到目前开源模型最高水平。Reflection 70B 的底层模型建立在 Meta 的 Llama 3.1 70B Instruct 上, 并使用原始的 Llama Chat 格式, 确保了与现有工具和 pipeline 的兼容性。

在通用能力之外, Reflection 70B 的亮点还包括“错误识别”和“错误纠正”。他们用到了一种名为“反思微调”的技术, 使得模型能够在

最终确定回复之前，先检测自身推理的错误并纠正。此外，它还将“规划”单独作为一个步骤，利用 CoT 缜密思考的过程，使得最终输出结果，更加简洁明了。

Reflection 70B 获得成功的另一个关键因素就是 Glaive 生成的合成数据。Glaive 的平台支持快速训练小型、高度集中的语言模型，有助于实现 AI 工具的普及。

Reflection 70B 的发布仅仅是 Reflection 系列的开端。而即将推出的 405B 版本，可谓是直接剑指诸如 GPT-4o 和 Claude 3.5 这些全球领先的专有或闭源 LLM。随着 AI 的发展，Reflection 在推理和错误校正方面的独特方法，很可能会成为开源模型的新标准。

AI 融资动向：SSI/智谱 AI 分别获 10 亿美元/数十亿元人民币融资

本期 AI 初创公司的融资中，SSI/智谱 AI 融资额前二，分别为 10 亿美元/数十亿元人民币。SSI 现拥有 10 名员工，SSI 明确表示构建安全超级智能是这个时代最重要的技术难题，他们已经启动了世界上第一个专注于 SSI 的实验室，目标和产品都只有一个：安全超级智能。智谱 AI 的 CEO 张鹏毕业于清华计算机系，以“让机器像人一样思考”为愿景，打造了对标国际领先水平的国产模型产品矩阵，包括 AI 提效助手智谱清言、高效率代码模型 CodeGeeX、多模态理解模型 CogVLM、生成图片模型 CogView 等。

投资建议

建议关注以 AI 为核心的龙头厂商科大讯飞（002230.SZ）、高速通信连接器业务或显著受益于 GB200 放量的鼎通科技（688668.SH）。

风险提示

1) AI 底层技术迭代速度不及预期。2) 政策监管及版权风险。3) AI 应用落地效果不及预期。4) 推荐公司业绩不及预期风险。

公司代码	名称	2024-09-08 股价	EPS			PE			投资评级
			2023	2024E	2025E	2023	2024E	2025E	
002230.SZ	科大讯飞	34.00	0.28	0.40	0.56	121.43	85.00	60.71	买入
002368.SZ	太极股份	14.24	0.79	1.01	1.28	18.03	14.10	11.13	买入
688668.SH	鼎通科技	33.99	0.67	1.04	1.41	50.73	32.68	24.11	买入

资料来源：Wind，华鑫证券研究

正文目录

1、 算力动态：算力租赁价格平稳，AMD RDNA4 GPU 架构显卡曝光.....	4
1.1、 数据跟踪：算力租赁价格平稳.....	4
1.2、 产业动态：英特尔发布酷睿 Ultra 200V 系列处理器，AI 算力大幅提升.....	4
2、 AI 应用动态：PERPLEXITY 访问量环比增长 7.48%，开源模型新黑马 REFLECTION 70B 发布.....	6
2.1、 流量跟踪：Perplexity 访问量环比增长 7.48%.....	6
2.2、 产业动态：开源模型新黑马 Reflection 70B 发布，使用自我纠错技术.....	7
3、 AI 融资动向：SSI/智谱 AI 分别获 10 亿美元/数十亿元人民币融资.....	8
4、 行情复盘.....	9
5、 投资建议.....	11
6、 风险提示.....	11

图表目录

图表 1：本周算力租赁情况.....	4
图表 2：Ultra 9 288V 功耗对比.....	5
图表 3：2024.8.14-2024.8.20AI 相关网站流量.....	6
图表 4：Reflection 70B 与其他模型性能比较.....	7
图表 5：本周 AI 初创公司的融资动态.....	8
图表 6：本周指数日涨跌幅.....	9
图表 7：本周 AI 算力指数内部涨跌幅度排名.....	9
图表 8：本周 AI 应用指数内部涨跌幅度排名.....	10
图表 9：重点关注公司及盈利预测.....	11

1、算力动态：算力租赁价格平稳，AMD RDNA4 GPU 架构显卡曝光

1.1、数据跟踪：算力租赁价格平稳

本周算力租赁价格环比持平。具体来看，显卡配置为 A100-40G 中，腾讯云 16 核+96G 价格为 28.64 元/时，阿里云 12 核+94GiB 价格为 31.58 元/时；显卡配置为 A100-80G 中，恒源云 13 核+128G 价格为 8.50 元/时；阿里云 16 核+125GiB 价格为 34.74 元/时；显卡配置为 A800-80G 中，恒源云 16 核+256G 算力紧张。

图表 1：本周算力租赁情况

显卡配置	CPU	内存	磁盘大小 (G)	平台名称	价格 (每小时)	价格环比上周
A100-40G	16	96	可自定, 额外收费	腾讯云	28.64/元	0.00%
	12 核	94G	可自定, 额外收费	阿里云	31.58/元	0.00%
A100-80G	13	128	系统盘: 20G 数据盘: 50GB	恒源云	8.50/元	-
	16 核	125G	可自定, 额外收费	阿里云	34.74/元	0.00%
A800-80G	16	256	系统盘: 20G 数据盘: 50GB	恒源云	-	-

资料来源：腾讯云，阿里云，恒源云，华鑫证券研究

1.2、产业动态：英特尔发布酷睿 Ultra 200V 系列处理器，AI 算力大幅提升

英特尔正式发布了酷睿 Ultra 200V 系列处理器，首批搭载这款处理器的笔记本电脑在 9 月 24 日起发售，现在已经开启预售，包括宏碁、华硕、戴尔、惠普、联想、LG、微星和三星在内的 80 款 PC 产品。

该处理器为有史以来最高效的 x86 处理器，对比前代 Meteor Lake，整体功耗降低达 50%；实现 x86 平台最佳电池续航，在生产力用例中续航长达 20 小时；每线程处理性能比前代 Meteor Lake 提高 3 倍，峰值性能提高 8 倍。

AI 算力大幅提升，CPU+GPU+NPU 算力突破 120TOPS，是 Meteor Lake 的 10 倍；NPU 数量从两颗增加到四颗，单独的 NPU 4 算力达 48TOPS，是 Meteor Lake 的 4 倍；首次亮相的 X(e)2 图形微架构，在移动图形性能方面带来平均 30%的性能提升

Ultra 200V 系列处理器采用全新的 P-Core Lion Cove 架构和 E-Core Skymont 架构，性能核与能效核的调度逻辑进行了改变，以更好地分配性能与功耗，同时集成了独显架构看齐的 Xe2 核显；计算模块采用台积电 N3B 工艺代工，平台控制模块采用 N6 工艺，首次将 LPDDR5x 内存与计算模块封装在一起，类似智能手机的 SoC 结构。

英特尔表示，在相同性能表现下，Ultra 9 288V 比高通 X Elite 的旗舰处理器 X1E-80-100 功耗低 40%，与苹果 M3 芯片旗鼓相当。

图表 2: Ultra 9 288V 功耗对比



资料来源: APPSO, 华鑫证券研究

2、 AI 应用动态：Perplexity 访问量环比增长 7.48%，开源模型新黑马 Reflection 70B 发布

2.1、 流量跟踪：Perplexity 访问量环比增长 7.48%

本期（2024.9.02-2024.9.06）AI 相关网站流量数据：访问量前三位分别为 ChatGPT（654.4M）、Bing（414.8M）和 Canva（166.2M）；访问量环比增速前三位分别为 Perplexity（7.48%）、QuillBot（5.48%）和 Kimi（3.59%）；平均停留时长前三位分别为 Discord（00:12:42）、Character.AI（00:12:40）和 NotionAI（00:09:08）；平均停留时长环比增速前三位分别为 QuillBot（1.24%）、Character.AI（0.93%）和 Bing（0.49%）。

图表 3：2024.8.14-2024.8.20AI 相关网站流量

应用	应用类型	归属公司	周平均访问量 (M)	访问量环比	平均停留时长	时长环比
ChatGPT	聊天机器人	OpenAI	654.4	3.25%	6:03	0.28%
Bing	搜索	微软	414.8	0.75%	6:52	0.49%
Discord	游戏社区	微软	130.8	0.15%	12:42	-0.65%
Canva	在线设计	Canva	166.2	1.96%	8:17	-0.20%
Github	代码托管	微软	111.9	-0.27%	6:45	0.00%
Gemini	聊天机器人	谷歌	62.52	-2.77%	4:57	-1.00%
Character.AI	聊天机器人	Character.AI	43.17	-3.08%	12:40	0.93%
NotionAI	文本/笔记	Notion	36.83	1.18%	9:08	0.00%
QuillBot	释义工具	QuillBot	15.98	5.48%	4:04	1.24%
Kimi	聊天机器人	MoonshotAI	5.457	3.59%	2:51	-2.84%
DeepL	翻译工具	DeepL	45.84	-1.12%	8:57	0.19%
文心一言	聊天机器人	百度	4.563	-1.68%	2:51	-0.58%
Perplexity	AI 搜索	Perplexity	14.95	7.48%	6:20	-2.31%

资料来源：similarweb, 华鑫证券研究

2.2、产业动态：开源模型新黑马 Reflection 70B 发布，使用自我纠错技术

当地时间 9 月 6 日，CEO of HyperWriteAI Matt Shumer 发布 Reflection 70B，模型表现达到目前开源模型最高水平。Reflection 70B 的底层模型建立在 Meta 的 Llama 3.1 70B Instruct 上，并使用原始的 Llama Chat 格式，确保了与现有工具和 pipeline 的兼容性，还使用了错误自我纠正技术——Reflection-Tuning。

在通用能力之外，Reflection 70B 的亮点还包括“错误识别”和“错误纠正”。他们用到了—种名为“反思微调”的技术，使得模型能够在最终确定回复之前，先检测自身推理的错误并纠正。此外，它还将“规划”单独作为一个步骤，利用 CoT 缜密思考的过程，使得最终输出结果，更加简洁明了。

Reflection 70B 获得成功的另一个关键因素就是 Glaive 生成的合成数据。Glaive 由荷兰工程师 Sahil Chaudhary 创立，专注于解决 AI 开发中最大的瓶颈之一：高质量、特定于任务的数据的可用性。Glaive 的平台支持快速训练小型、高度集中的语言模型，有助于实现 AI 工具的普及。

Reflection 70B 性能大幅超越 Llama 模型，并与当前顶尖 LLM 相抗衡。具体来说，Reflection 70B 在与顶级闭源模型（Claude 3.5 Sonnet, GPT-4o）比较中，表现出色。在 MMLU、MATH、IFEval、GSM8K 中测试基准上，都击败了 GPT-4o。更值得一提的是，仅凭 70B 参数彻底击败 405B 的 Llama 3.1，差距显而易见。

图表 4：Reflection 70B 与其他模型性能比较

Benchmark	Reflection 70B	Claude 3.5 Sonnet	Claude 3 Opus	GPT-4o	Gemini 1.5 Pro
GPQA	55.3% (0-shot Reflection)	59.4%* (0-shot CoT)	50.4% (0-shot CoT)	53.6% (0-shot CoT)	—
MMLU	89.9% (0-shot Reflection)	88.7%** (5-shot) 88.3% (0-shot CoT)	86.8% (5-shot) 85.7% (0-shot CoT)	88.7% (0-shot CoT)	85.9% (5-shot)
HumanEval	91% (0-shot Reflection)	92.0% (0-shot)	84.9% (0-shot)	90.2% (0-shot)	84.1% (0-shot)
MATH	79.7% (0-shot Reflection)	71.1% (0-shot CoT)	60.1% (0-shot CoT)	76.6% (0-shot CoT)	67.7% (4-shot)
GSM8K	99.2% (0-shot Reflection)	96.4% (0-shot CoT)	95.0% (0-shot CoT)	—	90.8% (11-shot)
IFEval	90.13% (0-shot Reflection)	88.0%	—	85.6%	—

资料来源：新智元，华鑫证券研究

Reflection 70B 的发布仅仅是 Reflection 系列的开端。而即将在下周推出的 405B 版本，可谓是直接剑指诸如 GPT-4o 和 Claude 3.5 这些全球领先的专有或闭源 LLM。随着 AI 的发展，Reflection 在推理和错误校正方面的独特方法，很可能会成为开源模型的新标准。

3、 AI 融资动向： SSI/智谱 AI 分别获 10 亿美元/数十亿元人民币融资

本期 AI 初创公司的融资中， SSI/智谱 AI 融资额前二， 分别为 10 亿美元/数十亿元人民币。 SSI 现拥有 10 名员工， SSI 明确表示构建安全超级智能是这个时代最重要的技术难题， 他们已经启动了世界上第一个专注于 SSI 的实验室， 目标和产品都只有一个： 安全超级智能。 智谱 AI 的 CEO 张鹏毕业于清华计算机系， 以“让机器像人一样思考”为愿景， 打造了对标国际领先水平的国产模型产品矩阵， 包括 AI 提效助手智谱清言、 高效率代码模型 CodeGeeX、 多模态理解模型 CogVLM、 生成图片模型 CogView 等。

图表 5： 本周 AI 初创公司的融资动态

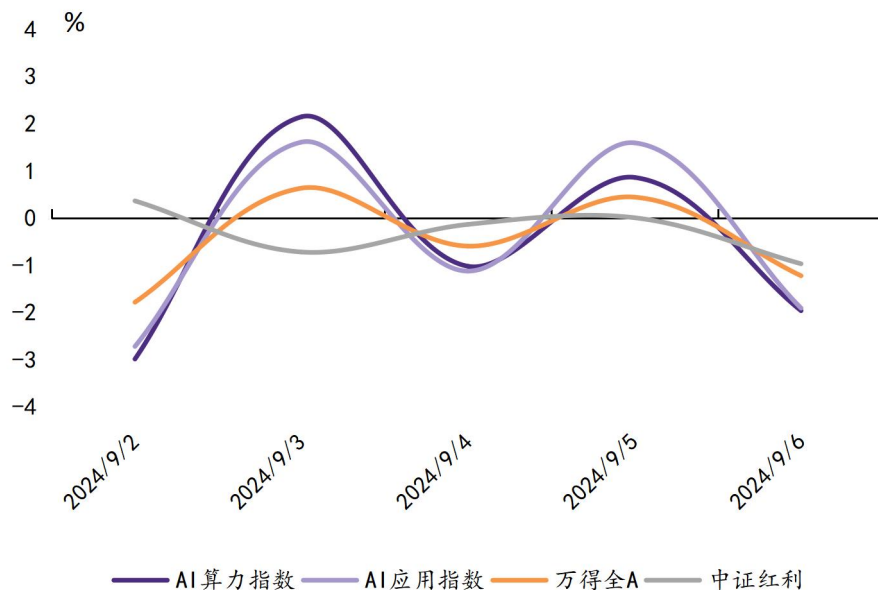
应用	应用类型	领投方	融资轮	融资额	目前累计融资额	目前估值
智谱 AI	AI 大模型	中关村科学城	——	数十亿元人民币	——	200 亿元
无问芯穹	AI Infra	君联资本、 启明创投、 洪泰基金	A 轮	5 亿元	10 亿元	——
SSI	安全超级智能	Andreessen Horowitz、 Sequoia Capital	——	10 亿美元	——	50 亿美元
Sakana AI	语言模型	Khosla Ventures	——	1 亿美元	——	——
Champion	AI 平台	Flyover Capital	种子轮	330 万美元	——	——

资料来源： 活报告， 企名片， 雷锋网， 澎湃新闻， 36Kr， 华鑫证券研究

4、行情复盘

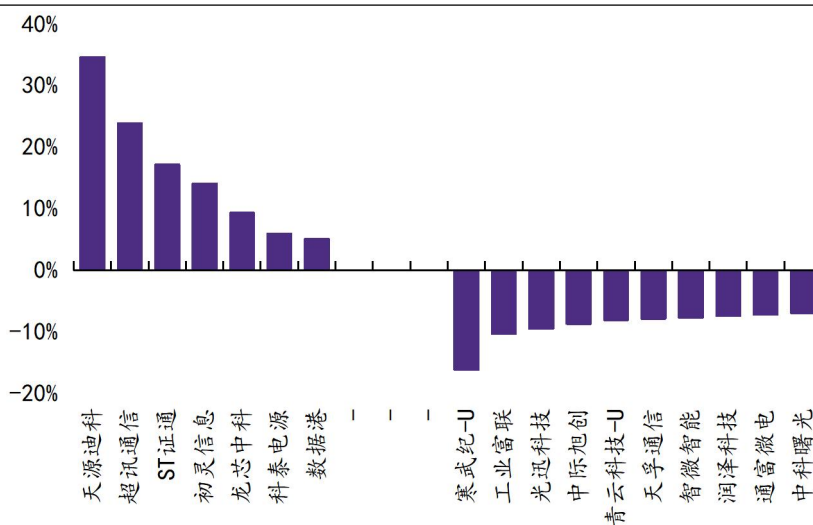
本周，AI 算力指数/AI 应用指数/万得全 A/中证红利日涨幅最大值分别为 2.15%/1.61%/0.64%/0.37%，日跌幅最大值分别为-2.98%/-2.72%/-1.78%/-0.96%。AI 算力指数内部，天源迪科以+34.67%录得本周最大涨幅，寒武纪以-16.17%录得本周最大跌幅。AI 应用指数内部，当虹科技以+12.68%录得本周最大涨幅，胜宏科技以-19.75%录得本周最大跌幅。

图表 6：本周指数日涨跌幅



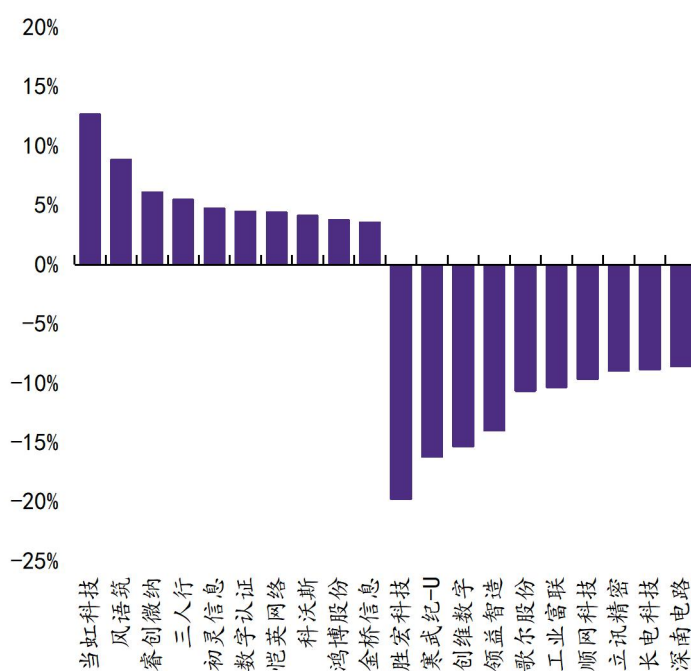
资料来源：wind, 华鑫证券研究

图表 7：本周 AI 算力指数内部涨跌幅度排名



资料来源：wind, 华鑫证券研究

图表 8：本周 AI 应用指数内部涨跌幅度排名



资料来源：wind, 华鑫证券研究

5、投资建议

对于近期市场对于 Blackwell 芯片设计存在问题的担忧，黄仁勋确认了 Blackwell 芯片将在四季度开始量产，并在四季度开始发货。

建议关注以 AI 为核心的龙头厂商科大讯飞（002230.SZ）、高速通信连接器业务或显著受益于 GB200 放量的鼎通科技（688668.SH）。

图表 9：重点关注公司及盈利预测

公司代码	名称	2024-09-08 股价	EPS			PE			投资评级
			2023	2024E	2025E	2023	2024E	2025E	
002230.SZ	科大讯飞	34.00	0.28	0.40	0.56	121.43	85.00	60.71	买入
002368.SZ	太极股份	14.24	0.79	1.01	1.28	18.03	14.10	11.13	买入
688668.SH	鼎通科技	33.99	0.67	1.04	1.41	50.73	32.68	24.11	买入

资料来源：wind, 华鑫证券研究

6、风险提示

1) AI 底层技术迭代速度不及预期。2) 政策监管及版权风险。3) AI 应用落地效果不及预期。4) 推荐公司业绩不及预期风险。

■ 计算机&中小盘组介绍

宝幼琛：本硕毕业于上海交通大学，多次新财富、水晶球最佳分析师团队成员，7年证券从业经验，2021年11月加盟华鑫证券研究所，目前主要负责计算机与中小盘行业上市公司研究。擅长领域包括：云计算、网络安全、人工智能、区块链等。

任春阳：华东师范大学经济学硕士，6年证券行业经验，2021年11月加盟华鑫证券研究所，从事计算机与中小盘行业上市公司研究

周文龙：澳大利亚莫纳什大学金融硕士

陶欣怡：毕业于上海交通大学，于2023年10月加入团队。

■ 证券分析师承诺

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并注册为证券分析师，以勤勉的职业态度，独立、客观地出具本报告。本报告清晰地反映了本人的研究观点。本人不曾因，不因，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收到任何形式的补偿。

■ 证券投资评级说明

股票投资评级说明：

	投资建议	预测个股相对同期证券市场代表性指数涨幅
1	买入	>20%
2	增持	10%—20%
3	中性	-10%—10%
4	卖出	<-10%

行业投资评级说明：

	投资建议	行业指数相对同期证券市场代表性指数涨幅
1	推荐	>10%
2	中性	-10%—10%
3	回避	<-10%

以报告日后的12个月内，预测个股或行业指数相对于相关证券市场主要指数的涨跌幅为标准。

相关证券市场代表性指数说明：A股市场以沪深300指数为基准；新三板市场以三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的）为基准；香港市场以恒生指数为基准；美国市场以道琼斯指数为基准。

■ 免责声明

华鑫证券有限责任公司（以下简称“华鑫证券”）具有中国证监会核准的证券投资咨询业务资格。本报告由华鑫证券制作，仅供华鑫证券的客户使用。本公司不会因接收人收到本报告而视其为客户。

本报告中的信息均来源于公开资料，华鑫证券研究部门及相关研究人员力求准确可靠，但对这些信息的准确性及完整性不作任何保证。我们已力求报告内容客观、公正，但报告中的信息与所表达的观点不构成所述证券买卖的出价或询价的依据，该等信息、意见并未考虑到获取本报告人员的具体投资目的、财务状况以及特定需求，在任何时候均不构成对任何人的个人推荐。投资者应当对本报告中的信息和意见进行独立评估，并应同时结合各自的投资目的、财务状况和特定需求，必要时就财务、法律、商业、税收等方面咨询专业顾问的意见。对依据或者使用本报告所造成的一切后果，华鑫证券及/或其关联人员均不承担任何法律责任。本公司或关联机构可能会持有报告中所提到的公司所发行的证券头寸并进行交易，还可能为这些公司提供或争取提供投资银行、财务顾问或者金融产品等服务。本公司在知晓范围内依法合规地履行披露。

本报告中的资料、意见、预测均只反映报告初次发布时的判断，可能会随时调整。该等意见、评估及预测无需通知即可随时更改。在不同时期，华鑫证券可能会发出与本报告所载意见、评估及预测不一致的研究报告。华鑫证券没有将此意见及建议向报告所有接收者进行更新的义务。

本报告版权仅为华鑫证券所有，未经华鑫证券书面授权，任何机构和个人不得以任何形式刊载、翻版、复制、发布、转发或引用本报告的任何部分。若华鑫证券以外的机构向其客户发放本报告，则由该机构独自为此发送行为负责，华鑫证券对此等行为不承担任何责任。本报告同时不构成华鑫证券向发送本报告的机构之客户提供的投资建议。如未经华鑫证券授权，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。华鑫证券将保留随时追究其法律责任的权利。请投资者慎重使用未经授权刊载或者转发的华鑫证券研究报告。