计算机

报告日期：2023 年 03 月 27 日

大模型如何影响网安行业的未来

——AI+网安专题报告

投资要点

- **我们认为大模型的本质是理解语言意图并根据意图进行任务分配，而一个行业是否具有语言体系以及流程性工作的占比是其能够被大模型赋能的关键标准。对于网络安全行业而言，网安产品具有较为标准的语言体系、且安服工作涉及大量流程性工作，故而使得大模型在网络安全行业的应用成为可能且值得期待。**

- **大模型的核心价值（以 GPT 为例）**

  GPT 的核心能力在于语义理解和文本生成。我们认为 GPT-3.5 引发市场轰动的原因在于其**相较于 GPT-3 突现能力的解锁，主要体现在复杂推理能力和泛化到新任务的能力**，从而使得其对于语义的理解能力有显著提升，**表现结果呈现与人类对齐。**

  同时我们认为以 **GPT 为代表的大模型的本质是理解语言意图并根据意图进行任务分配**，GPT-3.5 能力的突现使得其对于意图的理解更加精准,从而能够完成更加复杂场景的任务分配。

- **网安公司的主要业务**

  网络安全公司的业务类别可从产品、平台和服务三个层次来理解。进一步对各类任务进行抽象，网络安全防护主要涉及到四类任务：1） **管理网络安全语言体系；2） 阅读和生成网络安全语言体系，搜集整合信息；3） 理解网络安全语言体系；4） 基于对于意图的判定，执行对应反馈**。

- **大模型如何赋能网安行业**

  我们认为大模型不能够胜任的工作为网络安全语言体系的管理以及任务执行，**可以胜任的工作包括对于网络安全语言体系的阅读、理解、生成以及部分意图判定**。落地到具体应用场景而言，我们看好大模型赋能网安公司进行产品能力的提升以及安服团队的降本增效。

  1）产品能力提升：大模型可以帮助网络安全公司迭代现有产品，让产品更加智能，尤其是针对 Web 应用层的防护本身就涉及到对于语义的理解。

  2）安服团队降本增效：考虑到网安的日志为计算机领域的语言体系，大模型在 github 中预训练之后，对于日志的理解具有天然优势，可在安全运营中心（SOC）场景中降低安全服务人员的数量，实现降本增效。

- **投资建议**

  应用安全领域：安恒信息、绿盟科技、启明星辰

  日志审计领域：启明星辰、安恒信息

  安全取证领域：美亚柏科

  安全运维领域：启明星辰、安恒信息、奇安信、绿盟科技

  安全分析、情报、响应和编排（SAIRO）领域：奇安信、启明星辰、山石网科、绿盟科技、深信服、安恒信息

  安全服务领域：深信服、奇安信、安恒信息、启明星辰、、山石网科、亚信安全

- **风险提示**

  AI 技术落地不及预期；对于 AI 技术的投入或将导致行业内公司短期利润下滑；行业竞争加剧；宏观环境发生重大变化

行业评级：看好(维持)

分析师：刘雯蜀
执业证书号：s1230523020002
liuwenshu03@stocke.com.cn

研究助理：刘静一
liujingyi@stocke.com.cn

相关报告

1 《ChatGPT 插件发布，AI 应用生态加速构建》 2023.03.27

2 《数据要素行业双周报（二）——数据要素行业周报（2023.3.13-2023.3.24）》 2023.03.25

3 《2023 年计算机行业网络安全板块投资策略》 2023.03.22

正文目录

图表目录

1 从大模型的核心价值说起（以 GPT 为例）

自然语言理解（NLU）和自然语言生成（NLG）是自然语言处理（NLP）领域的两大核心任务，与之对应的，GPT 的核心能力亦体现在语义理解和文本生成的良好完成度上。伴随着多模态 GPT-4 模型的发布，GPT 的能力进一步从自然语言的理解和生成拓展至图像和文本领域。

图1： NLP 的两项核心任务：自然语言理解和自然语言生成

资料来源：企通查公众号、浙商证券研究所

在此基础上，我们延伸出两个问题进行探讨：

（1） 相较于此前的 GPT-3，GPT-3.5 模型的能力进步体现在哪里，从而引发了更加广泛的市场关注？

我们认为 GPT-3.5 引发市场轰动的原因在于其**相较于 GPT-3 突现能力的解锁，从而使得其对于语义的理解能力有显著提升，表现结果呈现与人类对齐**。根据安人心智董事长阳志平先生的测试，从人类理性思维的四类经典测试任务（语义错觉类、认知反射类、证伪选择类、心智程序类）对 GPT-3.5 进行评估，GPT-3.5 的正确率约为 58%，而人类测试正确率通常在 40%-60%之间，这意味着 GPT-3.5 能够生成更加符合人类期待的反馈（如零样本问答、生成安全和公正的对话回复、拒绝超出模型知识范围的问题）。

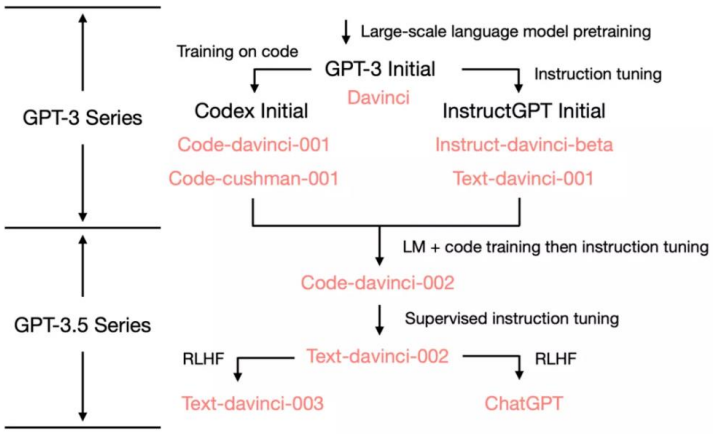
具体而言，**GPT-3.5 解锁的突现能力主要体现在复杂推理能力和泛化到新任务的能力**。根据艾伦人工智能研究所的研究，初代 GPT-3 主要展示**语言生成、上下文学习和世界知识**的能力，主要来源为大规模预训练；而 GPT-3.5 在 GPT-3 的基础上，进一步解锁了**响应人类指令、泛化到没有见过的任务、代码生成与理解、利用思维链进行复杂推理的能力**，主要通过代码训练、指令微调和基于人类反馈的强化学习(reinforcement learning with human feedback, RLHF)解锁。

表1： GPT-3 模型与 GPT-3.5 模型能力对比

GPT-3 模型		
	能力描述	**能力来源**
语言理解与生成	遵循提示词（prompt），然后生成补全提示词的句子	语言建模训练
上下文学习	遵循给定任务的几个示例，然后为新的测试用例生成解决方案	同一个任务的数据点在训练时按顺序排列在同一个 batch 中（推测）
世界知识	包括事实性知识(factual knowledge)和常识(commonsense)	3000 亿单词的训练语料库，而模型的 1750 亿参数是为了存储知识
GPT-3.5 模型		
	能力描述	**能力来源**
响应人类指令	GPT-3 的输出主要训练集中常见的句子。现在的模型会针对指令/提示词生成更合理的答案（而不是相关但无用的句子）	指令微调
泛化到没有见过的任务		指令微调：当用于调整模型的指令数量超过一定的规模时，模型就可以自动在从没见过的新指令上生成有效的回答
代码生成和代码理解		用代码训练模型
利用思维链进行复杂推理	参考人类解决问题的方法，从输入问题开始的一系列自然语言形式的推理过程，直到得到最后输出结论。初代 GPT-3 的模型思维链推理的能力很弱甚至没有。	代码训练的副产物（推测）

资料来源：新智元公众号、浙商证券研究所

图2： GPT-3 到 GPT-3.5 进化路径



资料来源：新智元公众号、浙商证券研究所

（2） 基于 GPT-3.5 甚至 GPT-4 的更加高级别的能力，在应用层面上提供了什么样的想象力空间？

我们认为以 **GPT 为代表的大模型的本质是理解语言意图并根据意图进行任务分配，从而实现对话、计算、制图等能力**。GPT-3.5 和 GPT-4 模型能力的实现使得其对于意图的理解更加精准，可**类比为人类世界中认知能力的提升**。在这个基础上，我们将 GPT 模型交互的流程框架粗略总结为 **“理解意图——拆解并分配任务——生成各项任务指令——通过接口分配至工具——以文本或图像形式返回结果”**，由于各项任务均可通过更加专业的工具完成，模型对于意图理解能力的大幅提升意味着其能够完成更加复杂场景的任务分配，可**类比为人类世界中调用工具完成任务能力的提升**。

图3： GPT 模型交互框架

资料来源：浙商证券研究所

2 网络安全公司在做什么

网络安全公司的业务类别可从产品、平台和服务三个层次来理解：

1）**产品：**网络安全产品分布在网络流量传输途径中的不同节点，基于预设的规则对网络数据及流量进行分析，并给出相应的任务反馈（拦截、放行、记录、告警等），其中各类流量和日志记录为网安产品进行各类操作记录的主要语言体系，规则标准相对固定。

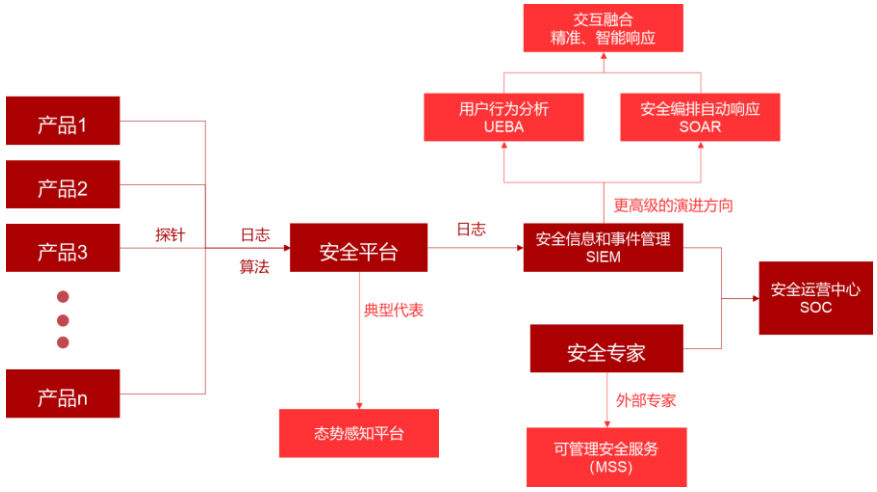
2）**平台：**网络安全平台可对多台网络安全产品进行统一管控，基于各产品自身的判定规则和任务标准，将多源信息汇总并基于更加全面、详细的规则进行深度分析，相较于产品而言，规则标准更加复杂。

以网络安全平台中的态势感知平台为例，基于分布于多个节点的流量探针所搜集到的数据，通过流量分析（利用元数据理解攻击活动）、特征分析（查找已知的恶意模式）、完整内容分析（借助全部数据包数据理解攻击活动）等方法实现检测、分析、响应、预测、预防、防御等功能。

3）**服务：**网络安全服务需求由安全运营中心（SOC）概念衍生而来。安全运营中心（SOC）是以资产为核心，以安全事件管理为关键流程，采用安全域划分的思想，建立一套实时的资产风险模型，协助管理员进行事件及风险分析、预警管理、应急响应的集中安全管理系统，强调**安全产品、平台、以及安全运维人员的结合**。

安全信息和事件管理（SIEM）是安全运营中心的核心组件，侧重于日志的集中式管理和审计，可以通过产品和平台完成交付，配合安全人员的人工干预，可实现安全日志的分析和安全风险的监控与定位，实现安全运营中心的目标。

图4：网络安全业务梳理



资料来源：中科网威网站、CSDN、安全牛公众号、浙商证券研究所

进一步对网络安全防护的各类任务进行抽象，我们认为主要涉及到四类任务：

1）**管理网络安全语言体系**：包含基本要素构建、预设数据和流量规则、将规则置入产品和平台体系；

2）**阅读和生成网络安全语言体系**，搜集整合信息：基于数据和流量特征，形成对应代码记录或日志；

3）**理解网络安全语言体系**：主要用来判断访问或攻击意图

4）**基于对于意图的判定，执行对应反馈**：执行拦截、放行、记录、告警等操作

当机器基于既定特征规则判定有异常情况或判定无法处理时，安全人员则会介入。

3 大模型能够帮助网安做什么

我们认为大模型不能够胜任的工作为网络安全语言体系的管理以及任务执行，可以胜任的工作包括对于网络安全语言体系的阅读、理解、生成以及部分意图判定。落地到具体应用场景而言，**我们看好大模型赋能网安公司进行产品能力的提升以及安服团队的降本增效**。

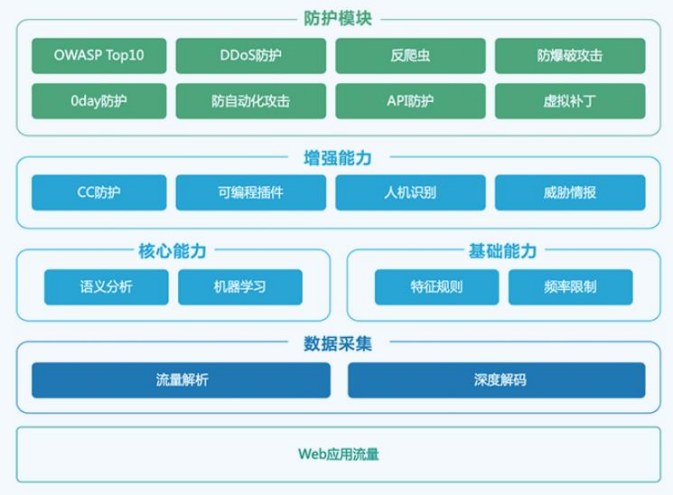
（1） 产品能力升级

大模型可以帮助网络安全公司迭代现有产品，让产品更加智能，尤其是针对 Web 应用层的防护本身就涉及到对于语义的理解。以 Web 应用防火墙（WAF）为例，传统 WAF 采用规则匹配的方式来识别和阻断攻击流量，随着 Web 漏洞数量的不断增加以及攻击方式不断变化，管理者通过持续更新和扩大规则库的方式来进行维护，维护工作繁琐且误报和漏报较为频繁。若赋予 WAF 更加高级的语义理解能力，就可以实现更加精准和智能的防护。

长亭科技的雷池（SafeLine）下一代 Web 应用防火墙即为全球范围内首款以智能语义分析算法为核心引擎能力打造的下一代 WAF，除形成了质变的检测引擎的精准程度外，还可以通过插件形式灵活扩展、实现瑞士军刀般的功能增加，同时可以变形适配、安装部署进各种网络环境，与机器学习等前沿技术更好的融合、增强流量分析的能力。

我们认为大模型对于语义理解能力的泛化推广或将对网安公司产生两方面影响：一方面，**率先布局的公司产品竞争力将大幅提升**；另一方面，**产品开发和运维效率有望得到大幅提升，助力网安公司研发团队以及下游客户实现降本增效**。根据长亭科技披露数据，其基于智能语义分析算法的下一代 WAF 产品可实现维护人员平均上线时间减少 50%，维护人员成本降低 70%。

图5：长亭科技基于智能语义理解的下一代 WAF 产品架构



资料来源：上海六有信息公众号、浙商证券研究所

（2） 安服团队降本增效

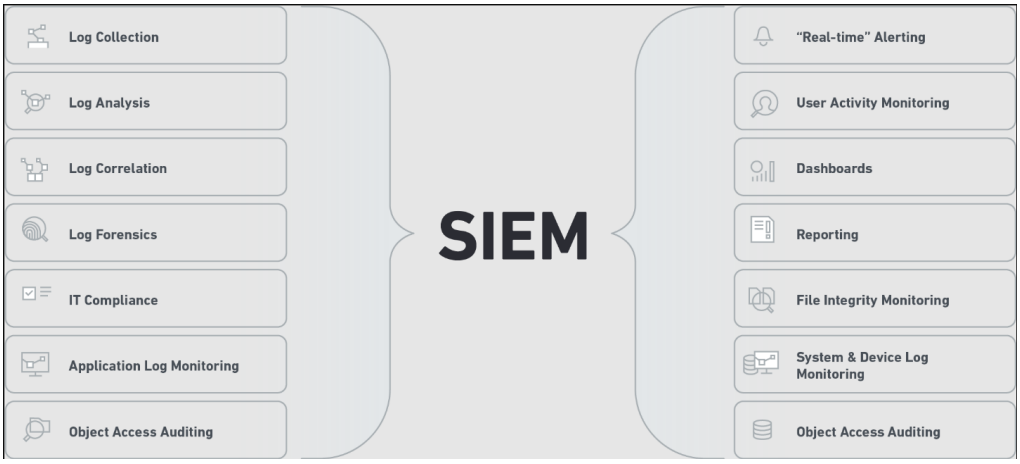
考虑到：1）网安的日志为计算机领域的语言体系；2）**GPT 在 github 中预训练之后，对于日志的理解具有天然优势**（基于无训练样本或少量训练样本的 zero-shot 和 few-shot learning 更具有应用的可能性），**我们认为大模型的应用有望在安全运营中心（SOC）场景中降低安全服务人员的数量，实现降本增效**。根据此前我们对于安全服务的描述，安全信息和事件管理（SIEM）是安全运营中心的基础，涉及到历史日志的分析和取证、实时分析日志和事件数据、提供威胁监控和时间响应。进一步来看，**更加高级别的 SIEM 已经发展到用户行为分析（UEBA）和安全编排自动响应（SOAR），其中已初步涉及大模型的语义理解和代码生成能力**：

- **用户行为分析（UEBA）**：传统的 SIEM 基于特征和规则进行分析，而用户行为超越了规则和相关性，故而可通过大模型的赋能研究人类行为模式，从而更加有效地检测内部威胁、针对性攻击和欺骗。

- **安全编排与自动响应（SOAR）**：涉及到将 SIEM 与企业系统集成并自动化响应事件。例如，在攻击者可以加密数据之前，SIEM 可能会检测到勒索软件的警报并在受影响的系统上自动执行应对操作，大模型的代码生成能力或将提升系统的自动响应能力。

安全服务是以企业资产、漏洞、威胁和事件四个要素作为核心抓手,通过安全专家团队和智能安全监测防护平台有效开展 7*24h 持续性网络安全保障工作,帮助企业单位大幅度降低发生安全事件的概率,同时较大幅度地提升整个企业组织的安全能力。**我们认为大模型具有大幅降低安全事件告警和响应频率的潜力，从而提升安全服务团队的人效，使得网安厂商的安服人员成本大幅降低。**

图6： SIEM 作用



资料来源：bbsmax 网站、浙商证券研究所

聚焦到具体场景而言，在重要活动安全保障期间一个典型的安全服务流程为“设备告警报错——前端安全人员排查日志，找出问题——遇到难以解决的问题，整理相关基础数据，上报高级别专家——溯源问题发生的原因，解决问题——维护设备规则库——记录事件日志，处置完结”，在这个流程过程中我们认为“前端安全人员排查日志”、“整理相关基础数据”、“维护设备规则库”、“记录事件日志”有望由大模型替代完成。

4 投资建议

通过对大模型核心能力以及其在网安领域的潜在应用场景进行梳理分析，**我们认为一个行业是否具有语言体系以及流程性工作的占比是其能够被大模型赋能的关键标准**。对于网络安全行业而言，网安产品具有较为标准的语言体系、且安服工作涉及大量流程性工作，故而使得大模型在网络安全行业的应用成为可能且值得期待。

从投资建议角度，我们建议从以下视角关注相关标的：

应用安全领域：安恒信息、绿盟科技、启明星辰

日志审计领域：启明星辰、安恒信息

安全取证领域：美亚柏科

安全运维领域：启明星辰、安恒信息、奇安信、绿盟科技

安全分析、情报、响应和编排（SAIRO）领域：奇安信、启明星辰、山石网科、绿盟科技、深信服、安恒信息

安全服务领域：深信服、奇安信、安恒信息、启明星辰、山石网科、亚信安全

5 风险提示

AI 技术落地不及预期：文章中关于 GPT 对于网安行业的赋能为结合行业现状和理论推理的假设思考，具体落地层面存在 AI 技术落地不及预期的风险；

对于 AI 技术的投入或将导致行业内公司短期利润下滑：由于在 AI 技术研发领域需要投入较多的人力和物理，或将拉高公司短期的费用率，导致利润下滑；

行业竞争加剧：单家公司的前置布局或将引起鲶鱼效应，导致其他公司竞相入场，从而导致行业竞争加剧；

宏观环境发生重大变化：宏观环境的大幅变化或将导致行业下游客户的预算分配及购买力发生变化。

股票投资评级说明

以报告日后的 6 个月内，证券相对于沪深 300 指数的涨跌幅为标准，定义如下：
1.买 入 ： 相对于沪深 300 指数表现 + 20％以上；
2.增 持 ： 相对于沪深 300 指数表现 + 10％ ~ + 20％；
3.中 性 ： 相对于沪深 300 指数表现 - 10％ ~ + 10％之间波动；
4.减 持 ： 相对于沪深 300 指数表现 - 10％以下。

行业的投资评级：

以报告日后的 6 个月内，行业指数相对于沪深 300 指数的涨跌幅为标准，定义如下：
1.看 好 ： 行业指数相对于沪深 300 指数表现 + 10%以上；
2.中 性 ： 行业指数相对于沪深 300 指数表现 - 10% ~ + 10%以上；
3.看 淡 ： 行业指数相对于沪深 300 指数表现 - 10%以下。
我们在此提醒您，不同证券研究机构采用不同的评级术语及评级标准。我们采用的是相对评级体系，表示投资的相对比重。
建议：投资者买入或者卖出证券的决定取决于个人的实际情况，比如当前的持仓结构以及其他需要考虑的因素。投资者不应仅仅依靠投资评级来推断结论。

法律声明及风险提示

浙商证券研究所

上海总部地址：杨高南路 729 号陆家嘴世纪金融广场 1 号楼 25 层
北京地址：北京市东城区朝阳门北大街 8 号富华大厦 E 座 4 层
深圳地址：广东省深圳市福田区广电金融中心 33 层
上海总部邮政编码：200127
上海总部电话：(8621) 80108518
上海总部传真：(8621) 80106010
浙商证券研究所：https://www.stocke.com.cn