

大语言模型在金融市场业务中的应用与探讨

◎宋云飞 岳一飞

摘要：近几年，以大语言模型为代表的 AIGC 模型在理论和实践上取得了长足发展，但在金融市场领域的应用仍处于早期探索阶段。本文对大语言模型的历史背景和发展历程进行了系统的回顾和梳理。结合金融市场业务的特殊性，本文总结了大语言模型在该领域的多个潜在应用场景、技术难点和业务价值。此外，本文对大语言模型在金融市场特定应用中存在的各种问题进行了归纳和探讨，结合我国实际国情，提出了在金融市场中应用大语言模型的技术性思考和政策性建议。本文认为，应加强大语言模型等技术在金融市场中的应用和探索，但需深入结合我国金融市场业务特性和风险管理要求，同时应完善流程规范和加强制度建设，这对于实现我国金融业数字化转型、高质量发展和高水平对外开放具有重要意义。

关键词：大语言模型；金融市场；AIGC；数字金融；智能风控

中图分类号：F832

文献标识码：A

党的二十大报告指出，“高质量发展是全面建设社会主义现代化国家的首要任务。”2023 年中央金融工作会议指出，“以金融高质量发展助力强国建设、民族复兴伟业”“加强优质金融服务，扩大金融高水平开放”“坚定不移走中国特色金融发展之路”。随着大数据、人工智能等技术在金融业的广泛应用，科技创新在金融高质量发展过程中任重道远。金融市场业务作为国内和国际金融市场的重要纽带，对构建国内国际双循环相互促进的新发展格局、推进我国金融高水平对外开放具有重要作用。近年来，大语言模型技术在多个领域取得了令人瞩目的成就，本文拟对该技术在金融市场中的应用及其对未来的战略意义进行探讨。

一、大语言模型的发展现状

（一）大语言模型的发展现状

自然语言处理是人工智能领域的经典问题，主

作者简介：宋云飞、岳一飞，中国银行全球市场部。

表 1 部分主流 GPT 相关模型的发布时间

模型名称	提出方	发布时间	参数数量 (亿)	预训练数据量	是否开源
Transformer	Google	2017-6	0.6-2	无	是
GPT-1	OpenAI	2018-6	1.5	5GB	否 (公开论文描述模型结构)
BERT	Google	2018-10	1-3	3.3 亿单词	是
XLNet	Google	2019	1-3	3.3 亿单词	是
GPT-2	OpenAI	2019-2	15-16	40GB	是
GPT-3	OpenAI	2020-6	1750	45TB	否 (可通过 API 访问模型, 下同)
GPT-3.5	OpenAI	2022-1	2000	-	否
ChatGPT (基于 GPT 的对话模型)	OpenAI	2022-12	-	-	否
GPT-4	OpenAI	2023-3	-	-	否

资料来源：作者根据公开资料整理

要研究如何使用机器算法去处理语言的两个基本功能——语义理解和语言生成。传统上，人们使用规则类模型或者统计模型等方法来实现一些简单的自然语言处理任务。近十年来，随着神经网络模型的深入研究和计算机硬件能力的提升，深度学习模型在自然语言处理问题中的应用取得了较大的进展。2017 年，Vaswani 等人提出 Transformer 模型，创造性地引入自注意力 (Self-Attention) 机制，实现了全局信息的交互和整合，显著提高了深度神经网络在文本领域的准确性。这是自然语言处理领域一个具有分水岭意义的成果。2018 年，Google 和 OpenAI 分别提出了更大规模的 BERT 模型和 GPT 模型。在此阶段之前，相关研究主要集中在模型本身，通过对仅编码 (Encoder-Only)，编码—解码 (Encoder-Decoder)，仅解码 (Decoder-Only) 等三种架构的探索来提高模型效果。2019 年开始，大语言模型呈现了爆发式的增长，模型研究方向也转向利用生成式预训练方法结合有监督微调的方式。2022 年，OpenAI 发布了 ChatGPT 应用，其在对话交流方面优秀的理解能力和表达能力获得了社会广泛关注。近年来部分主流的大语言模型的相关情况如表 1 所示。

大语言模型中比较有代表性的是 GPT (Generative Pre-Training) 模型。该模型从 Transformer 模型的基础上发展而来，两者结构基本类似，GPT 的模型结构如图 1 所示，由多个重复的解码器结构堆叠而成。GPT 模型首先引入了预训练 (Pre-Training) 方式，即无标签的数据集上进行生成式预训练，得到一个初级的通用语言模型；在此基础上，再针对特定目标任务，使用带标签数据集进行有监督微调 (Fine-Tuning) 得到准确性更高的特定

模型。这种预训练方式带来的主要好处是，可以在不影响准确率的前提下大幅提高模型的泛化能力，同时也扩大了可用训练数据的范围。传统的有监督训练方式需要对大量数据集进行手工标注，训练成本很高。而无监督的预训练方式减少了人工标注数据集的工作量，模型可以在海量无标注文本上进行训练，这使得模型规模的快速扩增成为可能。从 GPT-1 到 GPT-3，模型整体参数量由 1.17 亿左右增长到约 1750 亿，训练数据也从 5GB 增长到 45TB。

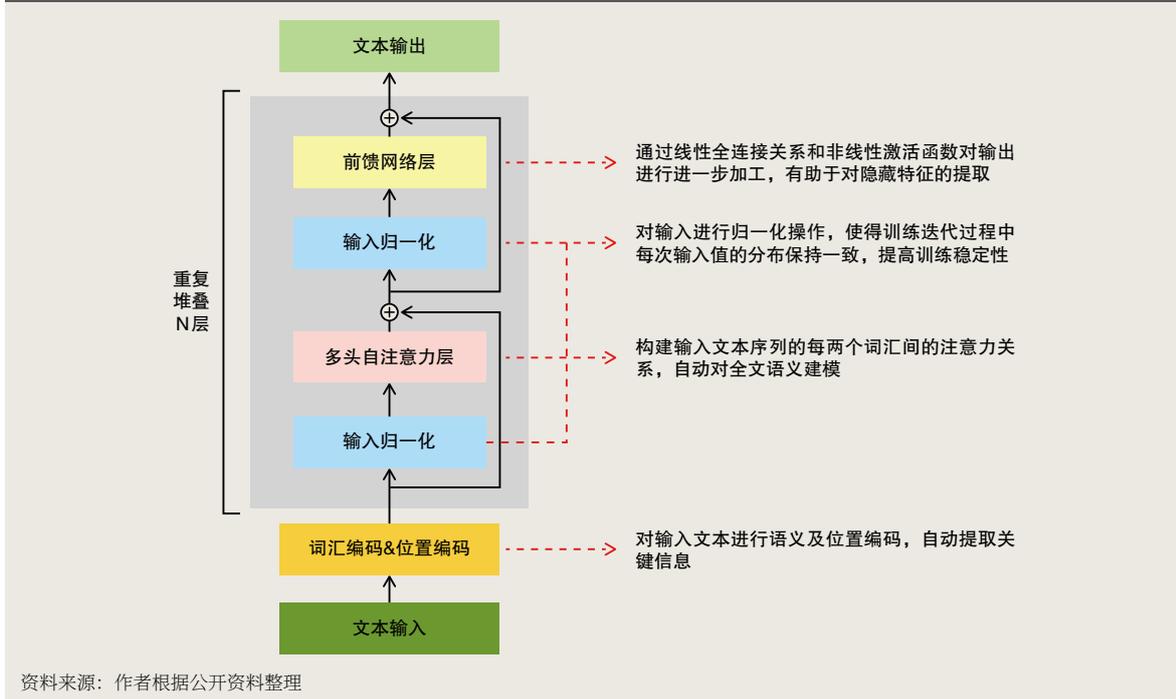
(二) 金融市场中的大语言模型概要

大语言模型 (Large Language Model, LLM) 是指通过包含大量参数的深度学习方法，解决自动翻译、模拟对话、文本内容生成等自然语言处理问题的模型。人工智能生成内容 (Artificial Intelligence Generated Content, AIGC) 模型泛指所有通过 AI 技术来生成新内容 (如语言、音频、视频等) 的超大规模模型。广义上来讲，大语言模型是 AIGC 模型的一个分支。近几年，大语言模型技术的实际应用日渐成熟，在多个领域取得了令人瞩目的成就。

大语言模型通常采用包含几百亿到几千亿个参数的深度神经网络模型，使用海量数据对模型进行训练。由于训练数据量巨大，大语言模型可以不针对特定的问题进行训练，而将训练后的模型用于有效解决多种文本处理相关问题。同时，由于对训练中使用的超大规模数据集进行人工数据标注几乎不可行，大语言模型通常采用自监督 (Self-supervised) 或半监督 (Semi-supervised) 的训练方式。

近些年，数字化转型是金融业的主要发展趋势

图1 GPT-2的模型结构图



之一。金融市场业务在金融体系中占据重要位置，为市场提供更好的流动性，为实体经济提供多维度的资金支持，为企业提供多种融资渠道和风控手段，是实现我国金融业高水平对外开放的重要组成部分。以大语言模型为代表的人工智能技术在金融市场业务数字化转型过程中有广阔的应用前景，对自然语言的准确理解、自动处理和精准分析是其重要环节。金融市场业务具有时效性高、容错率低的特点，因此在应用大语言模型技术时，对模型的安全性和可靠性提出了较高要求。由于市场上主流的大语言模型网络结构及参数未完全公开，同时又缺少针对金融市场的专业性，因此使用外部大语言模型处理金融市场业务具有一定风险。即使独立训练针对金融市场业务的大语言模型，也需要基于可靠的训练数据和训练过程，这点也存在一定的难度和挑战。

(三) 部分大语言模型产品现状

近年来，国内外多家公司和研究机构陆续推出了自己独立训练的大模型，这些模型可分为通用大模型和垂直领域大模型两种类型。

通用大模型是指不针对特定领域问题进行训

练，适用于解决各类任务的模型，该类模型的特点是规模都比较大。当前主流的模型有两种结构：GPT 类结构和 BERT 类结构。GPT 类结构主要采用生成（Generative）方式进行模型训练，其典型代表是 OpenAI 的 GPT 系列模型。GPT-3.5 是 ChatGPT 的基础模型，它可以理解自然语言的人工指令输入，与人进行对话以及完成各种自然语言处理相关任务。随后，ChatGPT 在新版本中增加了对图像等多模态数据的支持能力。其他 GPT 类结构还包括 BigScience 发布的 BLOOM 模型和 Meta 的 LLaMA 模型，二者均在生成式任务中有优秀表现。BERT 类结构主要采用判别（Discriminative）方式进行模型训练，其典型代表包括 Google 的 BERT 模型和 T5 模型等。国内相关公司也推出了类似的对话模型产品，例如，阿里云发布的“通义千问”和百度发布的“文心一言”等。清华大学发布并开源了 ChatGLM 模型，该模型采用 General Language Model (GLM) 架构，最高可达到 1300 亿参数，并在 1:1 比例的中英语料库上进行训练，旨在提高大模型对中文文本的处理能力。

垂直领域大模型是指针对特定领域任务，在训练中引入相关数据，以进一步提高模型在该领

域效果的大模型。此外，垂直领域大模型还可以通过外挂知识库方式进一步扩展模型的知识范围。例如，基于 LLaMA 模型微调并引入维基医疗数据的医疗领域大模型 ChatDoctor，基于 GPT 模型并引入 Github 代码数据进行微调的代码生成大模型 Codex，以及基于 LLaMA 模型训练并引入法律条文、司法解释、判决书等法律文本数据进行微调生成的法律领域大模型 ChatLaw 等。在金融领域，彭博公司在 2023 年 3 月发布了含有 500 亿个参数的 BloombergGPT 大模型。该模型在 BLOOM 模型基础上，引入彭博的大量金融新闻数据来进行训练，旨在提高针对财务金融领域文本的处理能力。

二、大语言模型在金融市场业务中的应用前景

（一）交易要素提取

金融市场的交易达成需要买卖双方达成协议来实现。现有的部分金融产品（如远期、掉期等），呈现市场流动性较差、自动化程度较低等特征。因此这类产品的交易达成模式，大多还是通过交易双方人工对话方式来完成。这一过程有较多人工参与的步骤，会占用交易员较多时间精力，且在交易活跃时段容易引发操作风险。如果对这部分交易进行自动化簿记，就需要从非标准的对话文本中自动化提取关键交易要素。利用大模型对自然语言的理解能力，可以实现这一目标。例如，可以使用脱敏处理后的交易对话数据对神经网络进行训练，得到大语言模型。通过对话达成交易后，可以将包含完整交易要素的对话文本输入大模型，大模型对输入的交易对话文本进行语义理解，从中提取标准化的关键交易要素，从而将其输入至下游系统，以完成后续自动化处理流程。相比传统的基于规则的方法，这种方法的优点是转换更加准确，适用于人工规则较难描述的输入场景，缺点是需要大量的交易对话数据来训练模型。

（二）文本逻辑分析

大模型在训练过程中可以获得对文本进行逻辑

分析的能力。对文本进行逻辑分析的主要场景包括内容概要生成、文章主题词生成、对文本中人物或事件之间的实体关系进行提取等。传统上，对文章关键信息的提取可以使用基于词频统计的 TF-IDF 方法或者基于文本中词汇连接关系的 TextRank 算法等。但这些方法在短文本等场景下都会面临结果不准确等问题。与这些方法相比，大模型可以从整个训练数据集中学习词汇、句子与文章之间的逻辑关系，掌握更加完整和复杂的语义信息，因此一般可以取得超过传统方法的效果。例如，大语言模型可以将新闻文本转化为只保留关键信息的内容概要，对市场信息的实时判断起到有效辅助作用，也可以将概要内容作为实时交易策略的输入特征，使策略模型有效利用新闻文本数据；对于主题词生成问题，大语言模型可以根据文本内容，判断文章主题以及关键词汇，在更广的范围内进行主题词标签选择和生成，弥补传统方法在部分场景下的不足；对于文本内容实体识别及实体关系抽取问题，在金融市场领域中，可以应用大语言模型，从自然语言表述的语义中抽取关键实体，并构建实体之间的相关关系，用于新闻文本的自动理解、规章制度的结构化表述等。

（三）智能交易对话

在通过对话达成交易的过程中，询价与报价双方需要采用文字方式交流交易意图及交易细节等，可能面临回复速度慢、容易出错等问题。通过大语言模型的文本生成能力，可以对询价意图及交易要素进行理解，自动化应答，完成交易对话。这种方法可以尽量真实地模拟人类的自然语言，一定程度上提高人工对话方式的交易效率。例如，从询价方角度，大模型根据指定的询价要求及对话内容等文本要素，可以自动生成具有逻辑关系的自然语言文本内容，通过程序接口进行自动交易对话；从报价方角度，可以接入大语言模型服务进行自动化对话报价。当报价方在收到人工询价时，通过从询价方的文字内容抽取交易要素等关键信息，再根据预设交易逻辑生成报价数据，并应用大模型将关键报价数据转化为报价文字内容，从而与交易对手进行自动化文本对话，以完成交易。

（四）金融知识图谱

大语言模型在训练过程中可以从训练数据中学习大量现有知识和逻辑关系等，并存储在模型参数中。由于训练数据中的信息以文本形式存储，大语言模型对这些文本信息的建模过程是以构建各个词汇之间相关性连接的方式完成的。相比于传统的结构化（如关系型数据模型等）方式，这种建模方式更为灵活，也更容易建立起包含多种关系类型的知识体系和内容关联逻辑。例如，大模型可以对金融市场相关的理论、实务知识进行建模，自动识别知识要点及其关联关系，形成具有多种关系连接的知识图谱网络，有助于对关键知识的多角度联想。也可以将金融领域的各层级规章制度等文本作为数据集，对大模型进行训练，作为交互式金融市场制度数据库，实现针对金融市场领域规则的交互式检索和联想，并可针对特定主题进行总结梳理，为实际工作提供便利。

（五）新闻事件信号

新闻和文本事件内容对金融资产价格的影响是市场广泛关心的驱动因素之一。市场关注的新闻焦点（如美联储会议纪要、联储主席讲话稿等）会对外汇价格形成实时而显著的影响。大语言模型由于能够同时处理大段文本内容输入，可以结合上下文信息对于同一词汇在不同语境中的不同含义进行精确区分，从而具有对文本内容的含义理解更准确的优点。例如，在实际应用时可以利用大语言模型对复杂文本的理解能力，将大语言模型进行进一步训练，将其作为一个基于文本内容进行价格变动方向预测的分类器，以文本内容、市场相关预测值等数据作为输入变量，进行文本情感倾向判断及价格变动方向的分类预测；也可以使用大模型结构对价格变动幅度进行预测，使其根据新闻等文本数据实时预测未来短时的价格变动幅度，从而争取合理获利机会，有效控制风险，辅助进行人工交易决策。

（六）智能风控

有效的风险管理对金融市场业务的发展至关重要。大语言模型可以对新闻文本、市场行情、经济指标等数据进行处理，从信用风险、流动性风险、

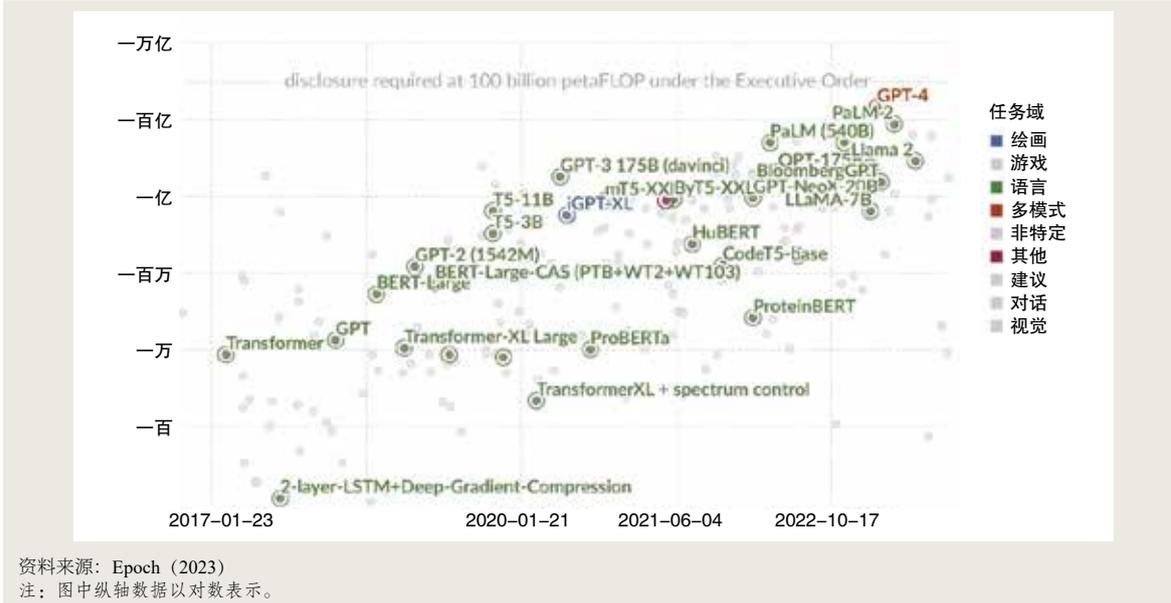
市场风险等多个角度进行辅助风险管控。由于大模型网络可以从多角度对大规模数据进行分析，大模型方法可以发现传统人工数据挖掘所难以发现的风险因子，可以提高风险控制的准确性和覆盖面，实现更加精准高效的风险管理。例如，大模型可以对历史违约事件相关新闻文本、市场行情等数据进行建模分析，建立信用风险监测预警模型，对交易对手的信用风险情况进行评估。当大模型在实时评估过程中碰到明显预期外的事件数据，如交易违约、负面舆论时，可以迅速针对相关主体进行风险提示，以减少交易对手违约等潜在风险损失。此外，大模型通过对大量跨市场历史行情数据进行统一建模处理，以市场实时行情作为输入，可以预测市场未来波动情况，及时识别异常模式，降低市场突然波动等潜在的市场风险。同时，大模型通过对操作记录数据的学习，可以识别正常交易行为的数据模式，在实际应用中自动识别潜在的异常交易行为并进行提示，降低整体操作风险。

三、大语言模型应用需要处理的“四个关系”

（一）“效率提升”和“风险可控”的关系

金融科技已经广泛应用到金融的各个领域，在可以预见的未来也将发挥更大的价值，并深刻地改变金融的生态。借助科技赋能，在数字化转型过程中，金融行业的效率将会得到大幅度的提升。然而应用新型技术存在面临新型未知风险的隐患，对于以防范风险为核心的金融体系来说需要格外谨慎。以大语言模型为代表的人工智能大模型在金融业的应用仍然处在初级阶段，其潜在价值需进一步发掘，潜在风险更需进一步衡量。单纯追求以新技术提升效率，可能会导致风险不可控；而单纯地追求降低风险，又有可能导致新技术难以应用而效率得不到提升。这也反映了科技和金融两种不同思维体系的差异。科技思维更注重实效、数据驱动，对大语言模型的准确性要求较高；而金融思维更注重底线思维、风险可控，对大语言模型的可解释性要求较高。因此，在金融市场领域应用大语言模型需要平衡好

图2 主流模型的训练成本呈现指数级增长



这两种思维的潜在风险和矛盾。

(二) “模型泛化”和“精准场景”的关系

截至2023年11月，国内外已经发布的大语言模型参数规模几乎都在百亿、千亿量级。这些大模型呈现了强大的对世界知识的掌握分析和对人类语言的理解能力，也展示出多模态能力。然而这些模型大多只开放给公众使用，而并没有把模型本身公开。对于一般性企业，训练此种规模的自研大语言模型有较大困难。为提高大模型技术在指定领域内的准确性，通常需要使用领域相关数据在通用模型基础上进行微调。但要使模型在领域内真正发挥作用，还需要对应用场景的深度理解。金融领域大模型不是简单地将金融数据引入到一般通用模型中进行微调而生成，而是需要对金融业务、金融数据、金融风险进行深刻理解，并将这些因素嵌入到模型应用场景中。避免出现“外行人看起来像内行，内行人看起来像外行”的情况。

(三) “数据私有”和“隐私保护”的关系

为提高在特定领域问题上的表现准确度，大语言模型在训练过程中需要使用大量针对该领域的相

关数据。对金融市场数据而言，数据中的敏感信息比例较高，如交易对手数据、客户信息数据等。为避免大模型在训练过程中保存敏感信息并出现在模型输出的内容中，需要对数据进行必要的脱敏、脱密等处理，同时仍要保持数据在模型训练中的有效性。此外，大模型在应用过程中一般采用在线学习（Online Learning）方式进行持续更新，在此过程中，用户使用模型时的交互记录、对话文本等数据也会进入模型训练过程中，这部分数据可能也包含用户敏感信息，也要对其进行敏感数据识别和脱敏等处理。

(四) “科技创新”和“管理缺失”的关系

以大语言模型为代表的一批新技术在金融市场中已经进行了广泛应用，为提高经营效率和加强风险管控提供了工具。在金融业高质量发展的背景下，将会有更多新技术应用于金融市场领域。但技术的快速发展不能脱离健全的金融监管。新技术的研究开发和落地应用过程中，可能会在技术伦理、隐私保护、风险防控、应用合法性等方面存在不足，需要在技术发展应用过程中不断加强制度建设，健全行业管理的体制机制。大语言模型技术具有强大的内容生成能力，可以生成现有文本范围以外的全

新内容，作为公共服务工具时会在社会中产生更复杂的影响，应当对其模型结构、数据来源及实际应用场景等进行有效管理。

四、大语言模型应用的思考与建议

（一）关于大语言模型应用于金融市场业务的思考

1. 大语言模型在金融市场业务中具有诸多优势
与传统的基于规则的方法相比，大模型对复杂数据的学习处理能力更强，适用问题所涉及数据量较大、人工进行规律总结较困难的情况。此外，当待解决问题的输入没有明确范围时，大模型在处理未知输入方面表现出了更大的优势，而决策树、随机森林等基于规则的方法以及传统机器学习领域的分类或回归模型方法在这种情况下可能较难发挥作用。然而，大语言模型在金融市场业务领域的广泛应用中仍面临许多实际困难。这些困难主要来源于：模型训练成本高、模型的不可解释性以及训练数据的代表性和隐私性问题。

2. 大语言模型对高质量训练数据需求量大，训练过程计算成本较高

首先，大语言模型对于特定任务的准确性和泛化能力高度依赖于训练数据的数量和质量。数据标注工作大多依靠手工完成，尽管模型预训练过程可以通过无监督学习方式，但在模型微调阶段，仍需要大量带标签数据进行训练。其次，无论是成交明细数据、交易过程对话数据，还是金融领域新闻等数据，都需要经过统一的数据校验和数据清洗，才能有效应用于模型的训练过程，数据清洗和维护工作将会耗费较多人力。最后，大语言模型的参数量巨大，训练过程需要消耗大量的计算资源和电力资源，经济成本较高。如图2所示，近几年主流大模型的训练成本呈现指数级增长。

3. 大语言模型在可解释性方面存在不足，需要就具体问题深入讨论

大语言模型采用深度神经网络，模型参数由数据集训练生成，无论对于学术界还是工业界，其解决实际问题的逻辑仍具有一定的不可解释性。因此，在金融市场业务具体实践中，仍需要人工确认

过程的参与。在理论方面，由于训练过程仅由机器根据训练数据完成，没有人工参与，因此模型自动习得的解决问题逻辑可能与人类的直观理解与判断逻辑不同，模型在一些情况下可能会出现看似简单的常识性错误，也可能产生符合语法规则但不符合事实的内容，即“模型幻觉”。在应用方面，当前大多数大型神经网络模型未完全开源，其模型细节、训练数据等也未在学术上得到充分的同行评审，在关键领域直接应用这些大模型产品的潜在风险仍然存在。在监管合规性方面，由于技术发展迅速，各国监管机构先后发布了人工智能技术在金融领域应用相关的规章制度，对模型可解释性提出了要求。2021年中国人民银行发布的《人工智能算法金融应用评价规范》中提出，要在算法建模准备、建模过程、建模应用等阶段对人工智能算法进行可解释性要求。

4. 大语言模型在数据代表性和数据隐私方面存在风险

大语言模型的能力与训练数据高度相关，模型经训练完成后，结构和逻辑即固定下来。但金融市场变化非常迅速，随着时间推移，模型训练所用数据可能不再能充分代表新的市场环境特征。若需要使模型适应新的市场情况，需要用包含新数据的数据集对模型进行重新训练，或者使用在线学习技术，定时使用新的数据对模型进行调整，提高模型在新数据集上的表现。大模型在数据隐私性管理方面也存在风险。模型训练过程中使用的大量数据可能包含未经授权的敏感数据，用户在使用大模型产品过程中输入的大量文本也可能包含用户隐私数据。如何对这部分数据进行管理，避免敏感数据泄露及滥用，对大模型服务提供者提出了挑战。

（二）大语言模型应用的启示与建议

1. 加强大语言模型在金融领域中的探索和应用，提升工作效率，更好服务实体经济

应用大语言模型等技术作为自然语言与机器语言之间的黏合剂，可以打破不同系统、流程之间的“自动化孤岛”现象，这有助于解决金融市场工作自动化、智能化过程中的瓶颈，大幅提高金融市场业务中的人力工作效率，提高金融市场运行效率。

高质量数据在模型训练和应用中具有重要意义。要加强数据积累,充分发挥现有数据的价值,适当增加新的数据来源,用优质的数据不断提高模型性能,帮助金融市场业务实现高质量发展。

2. 提升对大语言模型有效合理的风险管控,严守底线思维,防范金融风险

要客观看待大模型在金融市场业务数字化转型过程中扮演的角色。应用大语言模型技术过程应与适当的风险控制手段相结合,有效提高模型自身的风险控制能力和应对极端情况的能力,避免因模型的不可解释性而引发的潜在未知风险。要坚守底线思维,避免模型应用过程中出现极端不可控风险。要在模型训练和应用过程中保障数据隐私安全,避免数据泄露和滥用。同时,应当在可以实现目标的前提下尽量使用简单模型,避免模型结构过于复杂。

3. 推动金融机构对大语言模型技术的统筹管理,加强制度建设,提高治理能力

大模型的设计、开发、训练和应用过程,均会对模型效果和使用效率产生影响,也都可能出现数据隐私和技术伦理等问题。作为一种具有较高不可解释性特点的工具,大模型的全生命周期都需要纳入管理范围。开发和应用大模型的机构应当在训练数据、模型结构、应用场景等方面进行充分的技术论证和业务讨论。要制定健全的规章制度,保证在模型应用过程中的数据隐私安全,也要根据市场环境变化定时对模型进行更新,以确保模型效果,避免模型与预期不一致而导致的性能下降。

4. 完善大语言模型技术应用的合规监管,推动行业规范发展,保护投资者安全

目前,国内大模型研究处于上升阶段,高校、科研机构和企业均积极参与了大模型的研发,给产业、生活、教育带来了革命性发展。同时,为避免新技术衍生出的法律、道德、技术伦理问题,防止新行业出现野蛮生长,要在服务平台运营、内容管理、数据合规等多方面对大模型进行合规监管。2023年7月,国家网信办联合国家发展改革委等部门发布了《生成式人工智能服务管理暂行办法》,在数据及基础模型来源合法性、数据质量、知识产权管理、个人信息保护等方面明确了对大模型服务提供者的要求。未来,合规监管制度应继续进行深

化细化,保障行业健康发展,保护大模型使用者利益和信息安全。通过大模型技术促进金融高质量发展,开拓中国特色金融发展之路,是时代召唤、历史使命,也是落实金融服务实体经济根本宗旨的实际要求。

参考文献:

- [1] 常耀成、张宇翔、王红等. 特征驱动的关键词提取算法综述[J]. 软件学报, 2018, 29 (7): 2046-2070
- [2] 董胤蓬、苏航、朱军. 面向对抗样本的深度神经网络可解释性分析[J]. 自动化学报, 2022, 48 (1): 75-86
- [3] 秦兵、刘安安、刘挺. 无指导的中文开放式实体关系抽取[J]. 计算机研究与发展, 2015, 52 (5): 1029-1035
- [4] 杨涛. 对人工智能在金融领域应用的思考[J]. 国际金融, 2016 (12): 24-27
- [5] 于孝建、彭永喻. 人工智能在金融风险管理领域的应用及挑战[J]. 南方金融, 2017 (9): 70-74
- [6] Brown, Mann, Ryder, et al. Language Models Are Few-shot Learners[J]. Advances in Neural Information Processing Systems, 2020, 33: 1877-1901
- [7] Devlin, Chang, Lee, Toutanova K. Bert: Pre-training of Deep Bidirectional Transformers for Language Understanding[J]. Proceedings of naacL-HLT. 2019 (1)
- [8] Ji, Lee, Frieske, et al. Survey of Hallucination in Natural Language Generation[J]. ACM Computing Surveys, 2023, 55(12): 1-38
- [9] Radford, Narasimhan, Salimans, et al. Improving Language Understanding by Generative Pre-training[R]. 2018
- [10] Radford, Wu, Child, et al. Language Models Are Unsupervised Multitask Learners[R]. OpenAI blog, 2019, 1(8): 9
- [11] Vaswani, Shazeer, Parmar, et al. Attention Is all You Need[R]. Advances in Neural Information Processing Systems, 2017, 30
- [12] Wei, Tay, Bommasani, et al. Emergent Abilities of Large Language Models[J]. arXiv preprint arXiv:2206.07682, 2022

(责任编辑:辛本胜)