

基于大模型实现因果推断的探讨

Discussion on Causal Inference via Large Model

贾琳琳¹, 邓佳鑫¹, 庞俊彪¹, 张宝昌²(1. 北京工业大学, 北京 100124; 2. 北京航空航天大学, 北京 100191)

Jia Linlin¹, Deng Jiaxin¹, Pang Junbiao¹, Zhang Baochang²(1. Beijing University of Technology, Beijing 100124, China; 2. Beihang University, Beijing 100191, China)

摘要:

大模型是利用海量数据形成庞大参数数量的机器学习模型, 而因果推断是推断和理解事件、变量或行为之间的因果关系。从大模型与因果推断相结合的可能性和难点、大模型预训练、因果模型的人类反馈学习过程和微调技术等方面进行探讨, 论述了大模型具有发现潜在因果关系的机制和因果关系预测及解释的潜力。此外, 归纳了部分大模型开源工具, 可用于快速实现大模型训练、验证和部署。

关键词:

大模型; 因果推断; 机器学习

doi: 10.12045/j.issn.1007-3043.2023.07.004

文章编号: 1007-3043(2023)017-0020-05

中图分类号: TP181

文献标识码: A

开放科学(资源服务)标识码(OSID):



Abstract:

Large model is a machine learning model with huge parameters trained from massive data, and Causal inference is to infer and understand the causal relationship between events, variables or behaviors. It discusses the possibility and difficulty of the combination of the large model and Causal inference, the pre-training of the large model, the human feedback learning process of the causal model, and the fine-tuning technology. It shows that the large model has the mechanism of discovering potential causal relationships and the potential of causal relationship prediction and interpretation. In addition, it summarizes some open-source tools for large models, which can be used to quickly implement large model training, validation, and deployment.

Keywords:

Large model; Causal inference; Machine learning

引用格式: 贾琳琳, 邓佳鑫, 庞俊彪, 等. 基于大模型实现因果推断的探讨[J]. 邮电设计技术, 2023(7): 20-24.

1 概述

1.1 大模型

大模型指利用海量数据用自监督方法进行训练而生成拥有庞大参数数量的机器学习模型, 它们为下游任务提供了优良的分布式特征表示能力和模型泛化能力。例如, 处理复杂问答的多轮对话能力、人类情感对齐能力、数学问题的求解能力等。从大模型的发

展上看, 在 2022 年 11 月底, OpenAI 发布了 ChatGPT^[1], 2023 年 3 月 14 日, GPT-4 也随之发布。这 2 个模型的问世让全球切身感受到了人工智能从数据中形成智能的潜力。与此同时, 随着 MetaAI 开源了备受瞩目的 LLaMA^[2] 框架以及斯坦福大学提出了 Stanford Alpaca^[3] 模型, 业界开始涌现出更多类型的大模型。在国内, 百度、腾讯、华为、阿里等互联网头部企业竞相布局 AI 产业, 并推出“文心一言”“混元 AI 大模型”“盘古 Chat”“通义千问”等一系列通用化大模型。可见, 我国 AI 大模型发展加速, 行业的垂直应用也呈现出多样性、加

收稿日期: 2023-06-01

速性和广泛性。

1.2 因果推断

因果推断^[4-9]是一种统计和推理方法,旨在理解和推断事件、变量或行为之间的因果关系。因果推断可以被总结为3个层次。从问题角度出发,因果推断的第1个层次是相关性,即探索变量之间的关系。常规机器学习在应用中主要学习的就是变量之间的相关性。第2个层次是执因溯果,即研究实施某种干扰因素后产生的效果或者收益是否符合预期。第3个层次是由果执因,也被称为反事实推断,即为了得到某种结果需要怎么做。现在大多数因果建模都是从这一层次来进行探索。此外,多数因和果耦合在了一起,这可能导致因果推断无法获得满意的结果。因此,从因果推断的视角来看,将导致推断失误的因素解耦是一个至关重要的问题。与此同时,解耦因果同样对解决训练数据和测试数据的 Out-Of-Distribution (OOD) 问题起到帮助。它有助于更好地理解事件和变量之间的因果联系,从而支持更有效的决策制定和干预措施的设计。显然,因果推断在许多领域具有广泛的应用,包括公共政策评估、医学研究、经济学和社会科学等。

因果推断应用场景广泛,可以被用于医疗诊断和治疗决策、公共政策制定、金融风险评、自然灾害预测和防灾减灾等各个领域^[10-13]。具体如:因果推断在服务器问题诊断上的应用,而大模型可以分析服务器的运行数据,包括日志、传感器数据等。识别服务器中的异常行为和潜在问题,这使得预测服务器故障成为可能,并可及时采取相应的预防措施。同时,通过因果推断分析来确定问题的根本原因,并提供修复指导。

1.3 大模型与因果推断结合的可能性

大模型的强大分布式表示能力为因果推断提供了“因”和“果”的表示能力。即可以利用大模型来理解各种因果推断问题中的语义和表示,如因果发现和回答因果等问题。传统的因果发现和效应推理通常依赖于领域专家提供的先验领域知识,建立一些先验领域知识对相关领域的专家数量和知识互补性有很高的要求。然而,大模型的特征表示能力提供了先验领域知识的表示能力,支持自然语言和形式方法之间的转换。此外,大模型的泛化能力能够整合关于因果机制的常识和领域知识,从而因果关系在大模型的微调阶段、实践和采用方面开辟了可能性,有助于人

们更好地理解变量之间的因果联系。

因果推断对大模型的发展起到了促进作用。虽然,大模型能够在某些因果关系的简单问题上给出正确的答案,但目前的研究还不确定大模型是否能通过自监督学习掌握简单因果关系,并具备了因果推断的能力。然而,复杂因果关系及其可解释性注入到大模型学习的方法是还没有深入研究的领域。因此,复杂因果关系及其可解释性注入是大模型与因果推断结合的关键问题之一。相应地,基于大模型的因果推断的结论可解释性是大模型与因果推断结合的关键问题之二。

2 大模型因果推断的难点

大模型因果推断的主要难点在于实现从关联分析到因果推理的跨越。目前深度学习的主要特点是数据驱动、关联学习和概率输出,导致模型普遍存在预测不稳定和不可解释等问题。主要原因是因果机制尚未充分融入机器学习。一项研究^[14]探讨了大语言模型是否能够从相关性中推断出因果关系,并通过构建相关性推断因果关系的数据集 Corr2Cause 进行验证。该研究评估了17种现有的大语言模型在该数据集上的表现,结果显示目前的大模型在纯推理因果关系的能力方面存在不足,并且在这一任务上的表现基本上达到随机水平,并指出大模型的微调技术对因果关系推断具有重要意义。

因此,需要开展更多的研究以推动大模型在因果推断方面的发展,这些研究包括构建更适用于大模型因果推断的数据集,设计新的算法和模型结构以实现更好的因果推理能力,并提高模型在这一任务上的性能和稳定性。这将有助于推动大模型在因果推断领域的进展,并为机器学习带来更高水平的可解释性和可靠性。

3 大模型因果推断模型建立策略

3.1 大模型预训练

针对因果推断问题,大模型需要什么样的预训练策略和目标函数?首先,预训练模型是一种自监督学习的应用,利用几乎无限的文本(或行业数据),预测学习输入句子(或行业数据)中每个数据元(token)上下文相关的表示,它隐式地学习到了通用的语法语义知识(行业数据的变化规律)。第二,它可以将从开放领域学到的知识迁移到下游任务,以改善低标记资源

的下游任务。第三, 预训练模型+微调机制具备很好的可扩展性, 在支持一个新任务时, 只需要利用该任务的标注数据进行微调即可。第四, 需要构建体现因果推断的目标函数, 一种策略是基于外部因果陈述的自监督表示学习策略, 让自监督学习的特征表示能够识别外部定义的因果关系^[15-16]。

与经典的机器学习相比, 预训练模型具有三大特点^[17-19]。第一是模型参数规模大, 在预训练阶段充分利用大规模无标注数据, 使系统更好地掌握通用数据内在规律的能力^[20]。第二是泛化能力强, 在通用无标注数据上预训练得到的同一个大模型, 只需要对不同特定任务的有标注数据进行微调即可应用于不同任务中, 不需要针对每个任务专门研制模型。第三则是因其强大的通用语言能力而带来的良好的综合性能^[21]。

大模型进行预训练的一般步骤如下。

a) 收集大量低成本收集的训练数据。

b) 采用某种预训练方法去学习其中的共性, 一般采用自回归预训练方式或自编码预训练方式。

c) 将其中的共性“移植”到特定任务的模型中。

d) 预训练完成后, 可以进行迭代训练来进一步改进模型或压缩模型。这通常包括使用更具体的任务和数据集对模型进行微调, 以提高其在特定任务上的性能。

3.2 大模型因果推断的微调过程

首先, 微调可以提高模型在下游任务的性能。预训练的大型模型学习到了丰富的特征表示和模式识别能力, 通过微调, 可以利用这些学习到的特征表示快速、有效地适应新任务, 从而提高模型在特定任务上的性能。其次, 微调通常比从头训练模型更快, 起到加速训练过程的作用。预训练模型已经学习到了通用的特征表示, 因此微调只需要在特定任务的数据集上进行相对较少的迭代训练, 可以节省训练时间和计算资源。微调很利于迁移学习。通过微调大模型, 可以将在一个任务上学到的知识迁移到另一个相关任务上^[22-23]。

大规模语言模型在没有利用人类反馈技术(Reinforcement Learning from Human Feedback, RLHF)前, 已有的实验结果表明这些模型在多个数据集上的因果推断表现几乎接近随机性能^[14]。而通过人类反馈技术或因果推断标注数据集的微调后可以明显提升大模型的因果推断能力。具体地, 对于采用不同预训练

方式的模型采用不同的微调方法。例如, 基于GPT模型使用OpenAI调优API的默认设置; 而对于基于bert的模型, 基于transformer的自注意力机制适当交换bert的输入输出来进行微调。

微调一般包括以下步骤。

a) 加载预训练模型: 选择一个与所需任务相关的预训练模型, 并加载其权重。

b) 选择任务数据集: 选定特定任务所需的数据集, 目前已经有专门针对纯因果推断的数据集CORR2CAUSE。

c) 对模型进行微调: 将任务数据集作为输入, 以最小化模型在此数据集上的损失函数。在这个过程中, 通常需要在训练集和验证集上进行多次迭代, 以避免过拟合问题。

d) 在测试集上进行测试: 使用微调后的模型, 在测试集上测试其性能表现。

3.3 人类反馈技术提高因果推断的可解释性

OpenAI的研究者采用了人类反馈技术, 通过训练奖励函数对模型进行微调, 从而使其生成的内容更符合人类价值观^[24]。RLHF技术本质是强化学习, 在决策任务中, 让智能体的目标最大化累积回报。因此, 在涉及因果解释的场景中, 智能体的目标是最大化人类的满意度和因果推断结论的正确性。RLHF技术可以通过监督学习的方法学习到与人类反馈相一致的奖励函数, 从而可以将奖励与人类价值观和因果解释对齐。

在大模型中使用RLHF技术根据人类反馈进行训练时, 提示词(Prompt)被视为环境状态的一部分。模型将接收的提示词作为输入, 并根据该提示词生成输出, 再将输出展示给人类反馈者, 由其提供一个评价或指令来指导模型进行下一步操作。这个反馈信息可以被用作奖励信号, 用于调整模型的权重和参数, 使模型在输入上产生更好的输出。因此, 提示词在强化学习中扮演着重要的角色, 它用于指导模型的行为, 并且根据人类反馈进行优化。通过不断调整提示词和反馈过程, 可以提高模型的性能和生成质量。

提示词还是一种可以挖掘或促进大模型推理能力的技术思路。大模型本身是具备一些简单问题的推理能力, 但在复杂推理问题上需要提高。提示词设计的核心思想是通过合适的提示语或提示样本, 更好地激发出大模型本身就具备的推理能力。比如通过添加提示词“Let's think step by step”, 大模型在数学推

理任务上就可以进行多步推理并取得令人满意的效果。

在因果推断任务中,提示词一方面可以依据一些规则进行人工设计,设计提示词的一个有效方式是用思维链来使模型一步一步进行思考。一般设计规范是:<输入,思维链,输出>。其中,思维链允许模型将多步推理问题分解为中间步骤,逐步解决并给出最终答案。通常而言,提示词的设计分为以下3个步骤。

a) 模板设计。通过手动或自动设计模板,将输入 x (例如:独行月球真好看。)转变成 X (例如:独行月球真好看,这太___了),通常情况下 X 中包含空槽,让预训练语言模型对空槽填充,从而推断出 y (好看)。模板的设计灵活多变,需要根据下游任务与预训练语言模型来选择合适的模板。

b) 答案搜索。预训练语言模型在答案空间中进行搜索,找出得分最高的值填充到对应空槽中。

c) 答案映射。通过答案搜索得到空槽对应填充值后,部分任务的槽值为最终结果,部分任务的槽值需要进行转换,将槽值对应到最终的输出标签 y (好看)。另外还有一些通用技巧如使用明确的指令、最后重复一遍指令、添加语法如标点符号以及标题等。另一方面,提示词可以通过程序生成,通常称为提示词微调^[25]。目前已有的提示词微调技术有自动调节提示词方法^[26],它通过梯度优化自动从一系列候选词中生成对目标任务最佳的提示词;斯坦福大学的Percy Liang团队提出的Prefix-tuning方法^[27],它在Transformer的输入层中加入可学习的提示词嵌入;VPT(Visual prompt tuning)方法^[28],它在输入和中间层的节点中加入可调的提示词等。

4 大模型开源工具

在大模型发展的过程中产生了较多开源工具,它们提供了丰富的功能和资源,可以帮助研究人员和开发者更方便地构建、训练和部署大模型。它们在开源社区中被广泛使用和支持,并持续更新和改进,推动了大模型的发展和应用。

Transformer Reinforcement Learning (TRL)^[29]是用于用强化学习训练语言模型的库。通过TRL,可以用近似策略优化(Proximal Policy Optimization, PPO)训练Transformer语言模型。PPO是强化学习目前最有效的一种算法,和先前的强化学习算法相比,它在每一步迭代中都会尝试计算新的策略,可以让损失函数最小

化的同时保证与上一步迭代的策略间的偏差相对较小。该库建立在Hugging Face的Transformer库之上,目前已经支持GPT-2大模型进行训练。使用PPO算法优化GPT-2的大致流程可以分为续写(Generation)、评估(Evaluation)和优化(Optimization)3步。这可以使其在生成文本时更加符合预期的要求,并且尽量避免与参考语言模型的差异过大,有助于提高模型的性能和结果的质量。

Reinforcement Learning for Language Model (RL4LM)^[30]是Allen Institute for AI的一个用于语言模型微调和评估的开源库,它提供了多种强化学习算法(如PPO、NLPO、A2C和TRPO)、奖励函数和指标作为构建块。这个库的设计使得用户可以轻松地根据自己的需求进行定制,可以基于任何编码器-解码器或基于encoder transformer的语言模型进行训练,并使用任意用户指定的奖励函数。RL4LM的优势在于其灵活性和可定制性。它在实际任务中得到了验证,并计划支持更大模型和分布式训练,以进一步提升性能和扩展能力。

大模型开源工具大大促进了大模型的发展,通过共享知识、加速研究和开发、降低门槛和成本、社区合作和反馈以及创新和应用拓展,推动了大模型领域的快速发展和广泛应用。

5 总结

通过对大模型预训练、人类反馈学习过程和因果推断的微调技术等方面进行探讨,发现大模型在因果推断方面的巨大潜力。构建更适用于大模型因果推断的数据集,设计新的算法和模型结构可以实现更好的因果推断能力和推断的解释能力,提高大模型在因果推断上的性能和稳定性,促使大模型和因果推断发展形成良性循环。随着数据和模型不断发展,可以期待大模型在因果推断任务上应用能力将不断提升,并为各个领域理解和应用因果推断提供更多帮助。

参考文献:

- [1] BROWN T B, MANN B, RYDER N, et al. Language models are few-shot learners [C]//Proceedings of the 34th International Conference on Neural Information Processing Systems. Vancouver, BC, Canada: Curran Associates Inc., 2020: 1877-1901.
- [2] TOUVRON H, LAVRIL T, IZACARD G, et al. LLaMA: open and efficient foundation language models [DB/OL]. [2023-03-16]. <https://arxiv.org/abs/2302.13971>.

- [3] WANG Y Z, KORDI Y, MISHRA S, et al. Self-instruct: aligning language models with self-generated instructions [DB/OL]. [2023-03-20]. <https://arxiv.org/abs/2212.10560>.
- [4] 马忠贵, 徐晓晗, 刘雪儿. 因果推断三种分析框架及其应用综述 [J]. 工程科学学报, 2022, 44(7): 1231-1243.
- [5] 王东明, 陈都鑫. 因果推断: 起源和发展 [J]. 控制工程, 2022, 29(3): 464-473.
- [6] 赵慧, 郑忠国, 许静. 含隐变量和选择偏差的图模型中的因果推断 [J]. 北京大学学报: 自然科学版, 2006, 42(5): 6.
- [7] 金华, 方积乾. 强可忽略处理分配下因果推断的结构回归模型 [J]. 华南师范大学学报: 自然科学版, 2000(4): 6.
- [8] 张浩, 郝志峰, 蔡瑞初, 等. 基于互信息的适用于高维数据的因果推断算法 [J]. 计算机应用研究, 2015, 32(2): 4.
- [9] 凌昱玟. 复杂生产过程中若干因果推断问题研究 [D]. 上海: 上海交通大学, 2010.
- [10] 冯瑞玲, 郭茜. 教育领域的自然实验因果推断研究——以语言教育研究的应用为例 [J]. 清华大学教育研究, 2023, 44(2): 88-98, 148.
- [11] 马真真, 万亚平, 刘纯, 等. 基于因果推断肺癌患者生存时间预测方法 [J]. 计算机应用与软件, 2023, 40(4): 47-53.
- [12] 杨君, 田晨晨, 蒋墨冰, 等. 外资身份认定政策与企业创新策略——基于机器学习的因果推断 [J]. 国际贸易问题, 2022(10): 140-156.
- [13] 张晓阳. 基于大数据因果推断的大学生网购服装推荐模型研究 [D]. 哈尔滨: 哈尔滨工业大学, 2023.
- [14] JIN Z J, LIU J R, LYU Z H, et al. Can large language models infer causation from correlation? [DB/OL]. [2023-03-20]. <https://arxiv.org/abs/2306.05836>.
- [15] ZUO X Y, CAO P F, CHEN Y B, et al. Improving event causality identification via self-supervised representation learning on external causal statement [C]//Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021. Online: Association for Computational Linguistics, 2021: 2162-2172.
- [16] SONAWANE P, DROLIA S, SHAMSI S, et al. Self-Supervised Visual Representation Learning Using Lightweight Architectures [R/OL]. [2023-03-16]. <https://arxiv.org/abs/2110.11160>.
- [17] RADFORD A, NARASIMHAN K, SALIMANS T, et al. Improving language understanding by generative pre-training [R/OL]. [2023-03-16]. https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf.
- [18] DEVLIN J, CHANG M W, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding [C]//Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers). Minneapolis, Minnesota: Association for Computational Linguistics, 2019: 4171-4186.
- [19] HATAMIZADEH A, XU Z, YANG D, et al. UNetFormer: A Unified Vision Transformer Model and Pre-Training Framework for 3D Medical Image Segmentation [R/OL]. [2023-03-16]. <https://arxiv.org/abs/2204.00631v1>.
- [20] KAPLAN J, MCCANDLISH S, HENIGHAN T, et al. Scaling laws for neural language models [DB/OL]. [2023-03-16]. <https://arxiv.org/abs/2001.08361>.
- [21] WEI J, TAY Y, BOMMASANI R, et al. Emergent abilities of large language models [DB/OL]. [2023-03-16]. <https://arxiv.org/abs/2206.07682>.
- [22] LIALIN V, DESHPANDE V, RUMSHISKY A. Scaling down to scale up: a guide to parameter-efficient fine-tuning [DB/OL]. [2023-04-02]. <https://arxiv.org/abs/2303.15647>.
- [23] SCHICK T, SCHÜTZE H. Exploiting cloze-questions for few-shot text classification and natural language inference [C]//Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume. Online: Association for Computational Linguistics, 2021: 255-269.
- [24] LAMBERT N, CASTRICATO L, VON WERRA L, et al. Illustrating reinforcement learning from human feedback (RLHF) [EB/OL]. [2023-03-29]. <https://huggingface.co/blog/rlhf>.
- [25] LIU P F, YUAN W Z, FU J L, et al. Pre-train, prompt, and predict: a systematic survey of prompting methods in natural language processing [J]. ACM Computing Surveys, 2023, 55(9): 1-35.
- [26] SHIN T, RAZEGHI Y, LOGAN R L IV, et al. AutoPrompt: eliciting knowledge from language models with automatically generated prompts [C]//Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). Online: Association for Computational Linguistics, 2020: 4222-4235.
- [27] LI X L, LIANG P. Prefix-tuning: optimizing continuous prompts for generation [C]//Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers). Online: Association for Computational Linguistics, 2021: 4582-4597.
- [28] JIA M L, TANG L M, CHEN B C, et al. Visual prompt tuning [C]//Computer Vision - ECCV 2022. Cham: Springer, 2022: 709-727.
- [29] VON WERRA L, BELKADA Y, TUNSTALL L, et al. TRL: transformer reinforcement learning [EB/OL]. [2023-03-29]. <https://github.com/lvwerra/trl>.
- [30] RAMAMURTHY R, AMMANABROLU P, BRANTLEY K, et al. Is reinforcement learning (Not) for natural language processing: benchmarks, baselines, and building blocks for natural language policy optimization [DB/OL]. [2023-03-16]. <https://arxiv.org/abs/2210.01241>.

作者简介:

贾琳琳, 北京工业大学硕士在读, 主要研究方向为深度学习、神经网络等; 邓佳鑫, 北京工业大学博士在读, 主要研究方向为深度学习、持续学习等; 庞俊彪, 北京工业大学副教授, 博士, 主要研究方向为多媒体技术、计算机视觉、机器学习等; 张宝昌, 北京航空航天大学教授, 博士, 主要研究方向为复杂背景下快速小目标检测与识别研究、控制理论和深度学习的结合等。