AI 与数据隐私保护:联邦学习的破解之道

杨　强
(微众银行　广东深圳　518000)
(香港科技大学　香港)
(qyang@cse.ust.hk)

AI and Data Privacy Protection: The Way to Federated Learning

Yang Qiang
(*Webank*, *Shenzhen*, *Guangdong* 518000)
(*The Hong Kong University of Science and Technology*, *Hong Kong*)

Abstract　With the tremendous advance in computing, algorithms and data volume, artificial intelligence ushered in the third development climax, and began to gain a foot hold in exploring various industries. However, as the emergence of "big data", more "small data" or "poor-quality data", and "data silos" exist in industry applications. For example, in the information security realm, it is difficult for enterprises who provide security services such as content security auditing and intrusion detection based on artificial intelligence technology to exchange raw data due to the consideration of user privacy and trade secrets protection. The services between enterprises are independent, and the overall development of cooperation and technology is difficult to make a breakthrough in a short period of time. How to promote greater cooperation on the premise of protecting the privacy of organizations? Will there be any chance for technical means to solve the data privacy protection problems? Federated Learning is an effective way to solve this problem and achieve across-enterprise collaborative governance.

Key words　artificial intelligence; federated learning; data security; data privacy; corporate collaborative governance

摘　要　伴随着计算力、算法和数据量的巨大进步,人工智能迎来第 3 次发展高潮,开始了各行业的落地探索.然而,在"大数据"兴起的同时,更多行业应用领域中是"小数据"或者质量很差的数据,"数据孤岛"现象广泛存在.例如在信息安全领域的应用中,虽然多家企业推出了基于人工智能技术的内容安全审核、入侵检测等安全服务,但出于用户隐私和商业机密的考虑,企业之间很难进行原始数据的交换,各个企业之间服务是独立的,整体协作和技术水平很难在短时间内实现突破式发展.如何在保护各机构数据隐私的前提下促成更大范围的合作,能否通过技术手段破解数据隐私保护难题,联邦学习是解决这一问题、实现跨企业协同治理的有效方式.

关键词　人工智能;联邦学习;数据安全;数据隐私;企业协同治理

中图法分类号　TP309

收稿日期:2019-08-28

人工智能自1956年达特茅斯会议上被首次提出，已有60余年历史，但它真正能在商业上有所作为，能够高效化、规模化、普遍化地展现出社会经济潜力，则受益于过去10多年来计算力（云计算）、算法（深度学习等）和数据量（大数据）的巨大进步.从Google的AlphaGo开始，越来越多基于深度学习和神经网络的智能产品大量涌现并在应用领域崭露头角，科技巨头的加入大大增强了人工智能研究的资源和科研实力，人工智能的发展迎来第3次高潮，人工智能的研究从学者们个人的沙盘推演发展为大规模的团体作战.2016年AlphaGo总计使用了30万盘棋局作为训练数据并且接连战胜2位人类职业围棋选手，让大家看到了人工智能迸发出的巨大潜力，也更加憧憬人工智能技术可以在无人车、医疗、金融等更多、更复杂、更前沿的领域施展拳脚.技术的落地应用成为学术界和工业界共同关注的方向，更多的目光聚焦在新技术如何更好地转化为产业价值，“AI+行业”、“AI赋能行业”成为热门话题.

然而人工智能技术是否真的能如预想的一般在各行各业突飞猛进，当目光拉远到整体行业应用中时，问题随之而来——除了有限的几个行业，更多的应用领域有的只是小数据，或者质量很差的数据，并且这些数据分散在不同的机构，形成了一个个“数据孤岛”，虽然是参与规模庞大的团体作战，但缺乏有效的互通和协作.

以信息安全领域应用为例，虽然人工智能技术应用越来越广泛，比如基于神经网络，进行智能漏扫、入侵检测，实现主动安全防护和主动防御；在内容安全领域通过基于深度学习的OCR识别、敏感词检测等技术进行智能鉴黄、不良信息过滤，降低人工审核成本，净化互联网环境，目前也有多家企业推出了基于人工智能技术的安全内容服务.这些服务让更多机构，特别是以内容生产为核心的平台、网站在政策法规框架内创造更大价值.但目前各个企业之间的内容安全服务和数据都是独立的，在数据不能互通的情况下，各家的数据来源远远不够，模型效果很难大幅提升，要实现跨企业、政府、高校等多机构的整体安全协同治理也很困难.

那么能否把散落在各地、各机构的数据合并成大数据，这就存在着另一个问题——数据隐私保护.随着政策法规的逐渐完善和公众隐私保护意识的加强，如何在保护数据隐私的前提下实现行业协作与协同治理，如何破解“数据孤岛”与“数据隐私保护”的两难困境，成为当下人工智能技术行业应用中亟待解决的问题[1-3].

1 “数据孤岛”与“数据隐私保护”难题

数据孤岛和数据隐私保护的两难困境：一是来自于人工智能技术本身的特点，需要海量数据作为基础；二是来自于世界范围内对数据隐私和安全的日益重视.

人工智能技术尤其是深度学习依赖于模型、算法，更依赖于通过海量数据进行模型训练，从而不断改进，仅依靠某一机构所掌握的数据，无法实现技术的快速突破.理想状态是在数据之间建立广泛连接，形成合力，创造更大价值.而现实情况是：有效数据往往难以获取或以“数据孤岛”的形式呈现.公司之间的数据共享需要用户的授权，而许多用户倾向于拒绝数据共享；即便一个公司内部，数据壁垒也不易打通；互联网巨头的存在，使得少数公司垄断大量数据.这些因素都会导致数据孤岛，难以创造出“1+1>2”的数据价值.

全球范围内对数据隐私和安全的重视带来了更大挑战，这个挑战导致大部分企业只拥有小数据，加剧了数据孤岛现象的产生.欧盟出台了首个关于数据隐私保护的法案《通用数据保护条例》(General Data Protection Regulation，GDPR)，明确了对数据隐私保护的若干规定.和以往的行业规范不同，这是一个真正可以执行的法律，并且条款非常清晰严格.例如，经营者要允许用户来表达数据“被遗忘”的愿望，即“我不希望你记住我过去的数据，并希望从现在起，你不要利用我的数据来建模”.与此同时，违背GDPR的后果也非常严重，罚款可以高达被罚机构的全球营收的4%.Facebook和Google已经成为基于这个法案的第1批被告.而中国在2017年起实施的《中华人民共和国网络安全法》和《中华人民共和国民法总则》中也指出：“网络运营者不得泄露、篡改、毁坏其收集的个人信息，并且与第三方进行数据交易时需确保拟定的合同明确约定拟交易数据的范围和数据保护义务.”这意味着对于用户数据的收集必须公开、透明，企业、机构之间在没有用户授权的情况下不能

交换数据.

虽然有明确的法律法规并且在全球范围内达成了广泛共识,但由于技术等因素的限制,实际应用中,数据隐私保护仍然是难题.收集数据的一方往往不是使用数据的一方,如 A 方收集数据,转移到 B 方清洗,再转移到 C 方建模,最后将模型卖给 D 方使用.这种数据在实体间转移、交换和交易的形式违反了相关法律法规,并可能受到严厉的惩罚.

如何在保护数据隐私的前提下,从技术上解决数据孤岛的问题,在隐私、安全和监管要求下,如何让 AI 系统更加高效、准确地共同使用各自的数据,能够在小数据(很少的样本和特征)和弱监督(有很少的标注)的条件下做更好的模型,我们提出了联邦学习的解决方案,并且不断探索其在具体行业场景下的应用[4-8].

2 联邦学习——打破数据孤岛的必经之路

联邦学习(federated learning)指的是在满足隐私保护和数据安全的前提下,设计一个机器学习框架,使各个机构在不交换数据的情况下进行协作,提升机器学习的效果.其核心就是解决数据孤岛和数据隐私保护的问题,通过建立一个数据"联邦",让参与各方都获益,推动技术整体持续进步.

具体的实现策略是:建立一个虚拟的共有模型.这个虚拟模型类似于把数据聚合在一起建立的最优模型.但是在建立虚拟模型时数据本身不移动,因此不泄露隐私,符合数据合规要求,建好的模型也仅在各自的区域为本地的目标服务.在这样一个联邦机制下,各个参与者的身份和地位相同,实现"共同富裕".

微众银行 AI 团队联合社会各界,通过不断探索,提出了系统性的通用解决方案,并推动联邦学习技术在金融、医疗等行业落地应用,逐步建立联邦生态.

联邦学习有几大特征:

1) 各方数据都保留在本地,不泄露隐私也不违反法规;

2) 多个参与者联合数据建立虚拟的共有模型,实现各自的使用目的,共同获益;

3) 在联邦学习的体系下,各个参与者的身份和地位相同;

4) 联邦学习的建模效果类似于传统深度学习;

5) "联邦"就是数据联盟,不同的联邦有着不同的运算框架,服务于不同的运算目的.如金融行业和医疗行业就会形成不同的联盟.

在实际应用中,因为孤岛数据具有不同的分布特点,所以联邦学习也可分为:横向联邦学习、纵向联邦学习、联邦迁移学习 3 种方案,如图 1 所示:

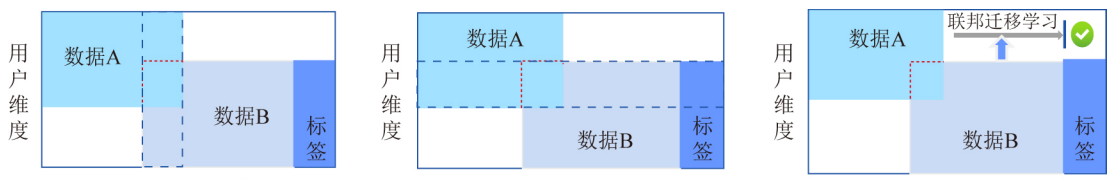


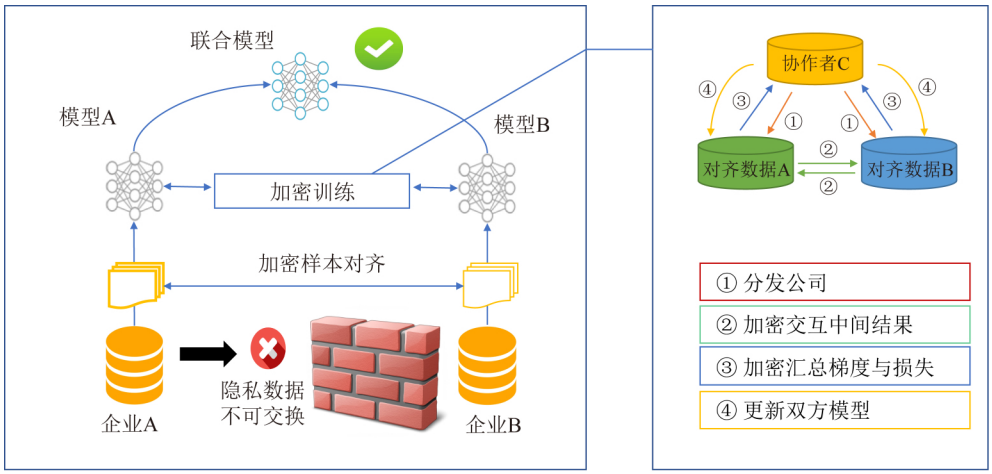
图 1 联邦学习的分类

如果要对用户行为建立预测模型,需要有一部分的特征,即原始特征,叫作 X,例如用户特征,也必须要有标签数据,即期望获得的答案,叫作 Y.比如,在金融领域,标签 Y 是需要被预测的用户信用;在营销领域,标签 Y 是用户的购买愿望;在教育领域,则是学生掌握知识的程度等.用户特征 X 加标签 Y 构成了完整的训练数据(X,Y).但是,在现实中,往往会遇到这种情况:各个数据集的用户不完全相同,或用户特征不完全相同.具体而言,以包含 2 个数据拥有方的联邦学习为例,数据分布可以分为 3 种情况:1) 2 个数据集的用户特征重叠部分较大,而用户重叠部分较小;2) 2 个数据集的用户重叠部分较大,而用户特征重叠部分较小;3) 2 个数据集的用户与用户特征重叠部分都比较小.为了应对以上 3 种数据分布情况,我们把联邦学习分为横向联邦学习、纵向联邦学习与联邦迁

移学习.

我们以包含2个数据拥有方(即企业A和企业B)的场景为例来介绍联邦学习的系统构架,该构架可扩展至包含多个数据拥有方的场景.

假设企业A和企业B想联合训练一个机器学习模型,它们的业务系统分别拥有各自用户的相关数据.此外,企业B还拥有模型需要预测的标签数据.出于数据隐私和安全考虑,企业A和企业B无法直接进行数据交换.此时,可使用联邦学习系统建立模型,系统构架由2部分构成,如图2(a)所示.



(a) 联邦学习系统架构组成　　(b) 加密模型训练

图2　联邦学习系统架构

1) 加密样本对齐:由于2家企业的用户群体并非完全重合,系统利用基于加密的用户样本对齐技术,在企业A和企业B不公开各自数据的前提下确认双方的共有用户,并且不暴露不互相重叠的用户,以便联合这些用户的特征进行建模.

2) 加密模型训练:在确定共有用户群体后,就可以利用这些数据训练机器学习模型.为了保证训练过程中数据的保密性,需要借助第三方协作者C进行加密训练.以线性回归模型为例,训练过程可分为以下4步(如图2(b)所示):①协作者C把公钥分发给模型A和模型B,用以对训练过程中需要交换的数据进行加密.②对齐数据A和对齐数据B之间以加密形式交互用于计算梯度的中间结果.③对齐数据A和对齐数据B分别基于加密的梯度值进行计算,同时对齐数据B根据其标签数据计算损失,并把这些结果汇总给协作者C.协作者C通过汇总结果计算总梯度并将其解密.④协作者C将解密后的梯度分别回传给模型A和模型B;模型A和模型B根据梯度更新各自模型的参数.

迭代上述步骤直至损失函数收敛,这样就完成了整个训练过程:在样本对齐及模型训练过程中,企业A和企业B各自的数据均保留在本地,且训练中的数据交互也不会导致数据隐私泄露.因此,双方在联邦学习的帮助下得以实现合作训练模型.

3) 效果激励.联邦学习的一大特点就是它解决了为什么不同机构要加入联邦共同建模的问题,即建立模型以后模型的效果会在实际应用中表现出来,并记录在永久数据记录机制(如区块链)上.提供数据多的机构会看到模型的效果也更好,这体现在对自己机构的贡献和对他人的贡献.这些模型会向各个机构反馈其在联邦机制上的效果,并继续激励更多机构加入这一数据联邦.

以上3个步骤的实施,既考虑了在多个机构间共同建模的隐私保护和效果,又考虑了如何奖励贡献数据多的机构,以一个共识机制来实现.所以,联邦学习是一个"闭环"的学习机制.

3　联邦学习的落地探索

目前,联邦学习已经开始了在行业领域的落地探索,在不同行业有多样化的应用场景和落地形态[9]:

在金融领域，多家机构联合建模的风控模型能更准确地识别信贷风险，联合反欺诈.多家银行建立的联邦反洗钱模型，能解决该领域样本少、数据质量低的问题.

在智慧零售领域，联邦学习能有效提升信息和资源匹配的效率.例如，银行拥有用户购买能力的特征，社交平台拥有用户个人偏好特征，电商平台则拥有产品特点的特征，传统的机器学习模型无法直接在异构数据上进行学习，联邦学习却能在保护三方数据隐私的基础上进行联合建模，为用户提供更精准的产品推荐等服务，从而打破数据壁垒，构建跨领域合作.

在医疗健康领域，联邦学习对于提升医疗行业协作水平更具有突出意义.在推进智慧医疗的过程中，病症、病理报告、检测结果等病人隐私数据常常分散在多家医院、诊所等跨区域、不同类型的医疗机构，联邦学习使机构间可以跨地域协作而数据不出本地，多方合作建立的预测模型能够更准确地预测癌症、基因疾病等疑难病.如果所有的医疗机构能建立一个联邦学习联盟，或许可以使人类的医疗卫生事业迈上一个全新的台阶.

目前，为了快速推进各行业联邦生态的建设，在工具层面，微众银行 AI 团队开源了首个工业级的联邦学习技术框架（federated AI technology enabler，FATE）[10-11]，不仅提供一系列开箱即用的联邦学习算法，更重要的是给开发者提供了实现联邦学习算法和系统的范本，使大部分传统算法可以经过改造适配到联邦学习框架中，从而快速加入联邦生态.

与此同时，相关的国际标准——IEEE 联邦学习标准制定也在推进中，今年 2 月份，IEEE P3652.1（联邦学习基础架构与应用）标准工作组第 1 次会议在深圳召开，目前国内外已经有 30 多个主要的企业和研究机构参与[12].作为国际上首个针对人工智能协同技术框架制定的标准，不仅会对各方联邦学习系统加以规范，还将为立法机构在涉及隐私保护的问题上提供技术参考.

联邦学习是一个大数据使用的新范式，是破解数据隐私保护难题的新思路，应用前景十分广阔.人工智能不仅是一个工具，更应该是让社会更加公平美好的强大推动力.我们或许已经掌握了一把钥匙，这扇大门却需要更多人一起推开.

参 考 文 献

[1] 中国计算机学会. 联邦学习助力 IoT？从“数据孤岛”走向“共同富裕”[OL]. [2019-08-15]. https://www.ccf.org.cn/c/2019-03-14/661020.shtml

[2] 黄善清. 不让“数据孤岛”成为 AI 发展的绊脚石，“联邦学习”将是突破口？[OL]. [2019-08-15]. https://www.leiphone.com/news/201902/5bLTrPeA6XwkwelR.html

[3] 丛末. AI 大数据在数据隐私保护下如何普惠共享？CCF TF「联邦学习」研讨会给出了答案[OL]. [2019-08-15]. https://www.leiphone.com/news/201903/qk0nnX5iC0G6bPaK.html

[4] Liu Y, Chen T, Yang Q. Secure federated transfer learning [J]. arXiv preprint, arXiv: 1812.03337, 2018

[5] Yang Q, Liu Y, Chen T, et al. Federated machine learning [J]. ACM Trans on Intelligent Systems and Technology, 2019, 10(2): 1-19

[6] Zhuo H H, Feng W, Xu Q, et al. Federated reinforcement learning [J]. arXiv preprint, arXiv: 1901.08277, 2019

[7] 杨强，刘洋，陈天健，等. 联邦学习[J]. 中国计算机学会通讯，2018，11(14)：49-55

[8] 杨强. GDPR 对 AI 的挑战和基于联邦迁移学习的对策[J]. 中国人工智能学会通讯，2018，8：1-8

[9] 微众银行 AI 项目组. 联邦学习白皮书 V1.0 [OL]. [2019-08-15]. https://www.fedai.org/static/flwp.pdf

[10] Cheng K, Fan T, Jin Y, et al. SecureBoost: A lossless federated learning framework [J]. arXiv preprint, arXiv: 1901.08755, 2019

[11] 微众银行 AI 项目组. 联邦学习开源平台 FATE [CP/OL]. [2019-08-15]. https://github.com/WeBankFinTech/FATE

[12] 木子. 微众银行 AI 团队领衔推动人工智能国际标准的制定[OL]. [2019-09-15]. https://www.leiphone.com/news/201902/BAmdBdnMrQSVQdXO.html

杨 强

博士，教授，香港科技大学新明工程学讲席教授，微众银行首席人工智能官.国际人工智能界迁移学习（transfer learning）技术的开创者，并提出联邦学习（federated learning）的研究新方向.于 2013 年 7 月当选为国际人工智能协会（AAAI）院士，是第 1 位获此殊荣的华人，并于 2016 年 5 月当选为 AAAI 执行委员会委员，是首位也是至今为止唯一的 AAAI 华人执委.2017 年 8 月当选为国际人工智能联合会（IJCAI）理事会主席，是第 1 位担任 IJCAI 理事会主席的华人科学家.主要研究方向为迁移学习、人工智能、大数据.

qyang@cse.ust.hk