联邦学习文献调研综述

22151203 刘龙鑫 软件工程

摘要

大数据的发展为数据驱动的机器学习和人工智能提供了强劲的动力。然而，随着数据利益的关注和隐私保护意识的不断提高，如何更加负责地使用人工智能成了人工智能研究社区亟需考虑的问题。联邦学习技术提供了一种数据可计算不可见的分布式模型构建方案，已在今年取得了长足的发展。本文就联邦学习的起源、定义，以及近年来的理论研究和实践应用情况作了简要综述。

keywords: 联邦学习、人工智能、数据安全、隐私保护

1 联邦学习概述

1.1 人工智能面临的挑战

最近几年，我们见证了人工智能的高速发展。计算机视觉、自然语言处理、语音识别等领域的研究引起了学术界和业界广泛热情。随着无数科学家和工程师的努力，许多科幻电影中的场景逐渐走进了我们的生活。目前各类人工智能模型取得的良好效果，绝大部分都是基于海量数据训练优化而来的。然而现实情况是，在大部分应用领域，训练出效果良好的模型所需要的数据量规模是难以甚至无法达到的。首先，采集整理并存储海量优质带标签的数据本身就是一件成本较高的任务；其次，随着基于人工智能模型提供的服务获得丰厚收益后，越来越多的组织和机构意识到了数据的价值，越来越重视数据的所有权，使得获取数据不再便利和廉价；另外，互联网时代的人们对于用户隐私和数据安全的重视程度也在不断提高。许多政府机构逐步推出新的数据保护法规，给数据的获取制造了更大的障碍，比如欧盟于2018 年 5 月 25 日出台的 GDPR（General Data Protection Regulation）[1]以及美国加利福尼亚州的 CCPA（California Consumer Privacy Act）[2]，对个人信息和侵害个人隐私的行为进行了明确且广泛的定义。我国的《中华人民共和国消费者权益保护法》、《中华人民共和国网络安全法》和《中华人民共和国民法通则》同样对非法侵害用户隐私行为提出了严格的管控要求。

如何在更加严格的隐私保护条例以及数据组织不愿再免费共享持有数据的情况下获取模型所需的训练数据，可能是当下人工智能模型落地所要面临的首要难题。为了解决这个难题，出现了一种由各个数据持有方在本地训练模型，然后通过一个精心设计的信息交换过程沟通并构建全局模型的方法，在这个信息交换过程中，任何一方都无法访问到其他数据持有方不愿泄露的隐私数据。这就是联邦学习的基本思想。

1.2 联邦学习的定义

H. Brendan McMahan 等人[3]于 2016 年率先提出了联邦学习的实用性概念。他们提出了一种中心化模型学习的替代方案，通过在存储有本地用户数据的客户端训练模型，并由服务器聚合计算更新来构建共享模型的方法来实现去中心化模型训练，在降低通信成本并在一定程度上保护用户隐私的前提下，为客户端应用提供更好的服务。

一般来说，联邦学习旨在基于一个分布式的数据集训练模型，模型的训练过程大多在数据持有节点本地完成，而模型相关信息则以加密的形式在各个节点之间交换。使得整个系统能够在保护各方数据隐私部分的同时，有效利用数据构建共享的学习模型。

1.3 联邦学习的分类

由于联邦学习数据来源的多样性，根据参与方数据特征空间和样本空间的分布情况，将联邦学习分成了三种基本形式：

（1） 横向联邦学习（Horizontal Federated Learning）：对于数据提供方数据特征重叠较多，而样本空间重叠较少的情况，适合采用按样本划分的联邦学习，也称作特征对齐的联邦学习（Feature-Aligned Federated Learning）；

（2） 纵向联邦学习（Vertical Federated Learning）：对于数据样本空间重叠较多，而特征空间重叠较少的情况，通常联合多个数据源的共同样本进行纵向的特征维度扩展，这种学习形式也称作样本对齐的联邦学习（Sample-Aligned Federated Learning）；

（3） 联邦迁移学习（Federated Transfer Learning）：这种联邦学习形式适用于数据提供方样本空间和特征空间重叠都较少的情况。需要挖掘领域之间的相似性。

1.4 隐私保护技术

联邦学习实现模型信息交流的设计要求中最重要的就是隐私信息的保护，这也是联邦学习的意义所在。对于隐私保护机制的设计，有许多隐私保护技术可以借鉴。

（1） 安全多方计算（Secure Multi-Party Computation）[4]：由华人计算机科学家姚期智于 1982 年提出。最初来源于一个安全两方计算问题，“百万富翁问题”。其目的是为了解决协同方在不泄露隐私输入的情况下，安全地计算出一个约定函数结果的问题。一般来说，安全多方计算有三种主流的实现框架：不经意传输（Oblivious Transfer）[5]、秘密共享（Secret Sharing）[6]和阈值同态加密（Threshold Homomorphic Encryption）[7]；

（2） 同态加密（Homomorphic Encryption）[8]：最早由 Rivest 等人于 1978 年提出。指的是一种能够进行特定运算的加密函数，可以实现对明文进行环上的加法或乘法运算后结果的加密，与加密后的相应运算结果等价。以作为一种不需要对密文进行解密的隐私计算解决方案。同态加密又分为部分同态加密（Partially Homomorphic Encryption）、些许同态加密（Somewhat Encryption）和全同态加密（Fully Homomorphic Encryption）；

（3） 差分隐私（Differential Privacy）：差分隐私是一种隐私概念，最早由 Dwork 于 2006 年提出，是一种常被用于抵抗成员推理攻击的具有随机性质的数据处理技

术。主要通过引入噪声的方式为意思数据提供差分隐私保护。主要有两种噪声选择方式，基于函数敏感性的噪声[9]和基于离散值指数分布的噪声[10]。

2 联邦学习的研究

在计算机科学与人工智能的发展历程中，联邦学习曾以不同的形式出现过。如分布式机器学习（Distributed Machine Learning）、面向隐私保护的机器学习（Privacy-Preserving Machine Learning）等。谷歌于 2016 年发表的论文 Communication-Efficient Learning of Deep Networks from Decentralized Data[3] 较为系统地给出了联邦学习的概念与应用框架。并在 Federated Learning for Mobile Keyboard Prediction[11] 中详细阐述了 FederatedAveraging 算法在 GBoard 用户输入预测中的应用。

主流的联邦学习研究工作主要着眼于隐私保护以及模型聚合上的工作。K Cheng 等人[12]提出了一种称为无损隐私保护树提升系统的联邦学习框架，并声称能够提供与非隐私保护技术相同级别的学习准确性。也有学者对于联邦学习系统中的恶意攻击进行了研究，AN Bhagoji 等人[13]在实验中通过交替最小化策略提高隐藏性，指出了即使是高度受限的恶意节点也能够在维持高度隐藏性的同时对模型发起投毒攻击。

对于联邦学习中不同机器学习模型的比较，也成为了一种新的研究方向。HH Zhuo 等人[14]在联邦学习的环境下提出了一种新的强化学习方法，利用共享信息的高斯差来实现模型参数的聚合，以保障隐私数据的安全性。

基于联邦学习分布式并发的特征，Virginia Smith[15]等人在引入高通信成本、信息延迟以及容错性的考虑的同时，对多任务学习在联邦学习框架下的性能表现以及优化措施进行了研究。

计算机科学家与工程师对于联邦学习的研究兴趣并非踱步于理论阶段，许多实际应用领域也逐渐见到了联邦学习的身影，如联邦学习框架下的计算机视觉任务在医学图像分析中的应用，自然语言处理在社交平台情感分析中的应用，基于隐私保护的推荐算法等等。此外，许多联邦学习的开源项目也处于迅速发展壮大阶段，如微众银行的 FATE（Federated AI Technology Enabler）[16]框架。

3 联邦学习的应用

3.1 金融领域

联邦学习已经在许多金融行业的业务中取得了一定进展，如联合信贷风控建模、信用卡诈骗检测等。W Yang 等人[17]基于银行本地数据库构建局部欺诈检测模型，通过聚合方法构建共享的 FDS 模型，并使用过采样方法平衡偏斜的数据集，给出的实验数据得到了优于传统 FDS 模型的检测效果。G Shingi 等人[18]同样引入了过采样技术来解决数据不平衡的问题，在联邦学习框架下为贷款违约风险评估构建了模型。

3.2 医疗领域

医疗行业相关的数据往往具有较高的分散性和敏感性，而联邦学习的介入为联合医学模型训练以及个性化医护方案提供了基础性推进作用。数据驱动的机器学习已成为一种有前途的方法，可以从现代医疗保健系统大量收集的医疗数据中构建准确而稳健的统计模型。模型没有充分利用现有的医疗数据，主要是因为它位于数据孤岛中，隐私问题限制了对这些数据的访问。 但是，如果无法获得足够的数据，机器学习将无法充分发挥其潜力，并最终无法从研究过渡到临床实践。Rieke 等人[19]考虑了导致该问题的关键因素，探讨了联邦学习如何为数字健康的未来提供解决方案，并强调需要解决的挑战和注意事项。

3.3 智慧城市

同样针对数据分散性和敏感性的痛点，联邦学习为智慧城市的构建提供了有力的数据支持。许多关于交通流量的预测研究取得了可观的进展。Yi Liu 等人[20]意识到了当代城市居民、出租车司机、商业部门和政府机构对准确及时的交通流量信息有着强烈的需求。人们可以使用交通流量信息来制定更好的出行计划。交通流预测（Traffic Flow Prediction，TFP）就是利用历史交通流数据来提供这样的交通流信息来预测未来的交通流。 TFP 被认为是成功部署智能交通系统 (ITS) 子系统的关键要素，尤其是先进的旅行者信息、在线叫车和交通管理系统。

4 总结

诞生于数据碎片化、数据孤岛、隐私泄露问题日益受到关注的互联网高速发展时代，联邦学习在试图构建健全完善的隐私数据保护生态系统的同时，为数据驱动的机器学习、人工智能模型提供了重要的数据源解决方案。然而许多理论与实践仍然有很大的改进发展空间，除了在隐私保护的前提下提供更好的智能化服务，如何构建一个公平透明的共享激励生态系统还是许多联邦学习系统能够进一步发展的重要课题。

5 参考文献

[1] Voigt, Paul, and Axel Von dem Bussche. "The eu general data protection regulation (gdpr)." A Practical Guide, 1st Ed., Cham: Springer International Publishing 10 (2017): 3152676.

[2] Pardau, Stuart L. "The California consumer privacy act: towards a European-style privacy regime in the United States." J. Tech. L. & Pol'y 23 (2018): 68.

[3] McMahan, Brendan, et al. "Communication-efficient learning of deep networks from decentralized data." Artificial intelligence and statistics. PMLR, 2017.

[4] Yao, Andrew C. "Protocols for secure computations." 23rd annual symposium on foundations of computer science (sfcs 1982). IEEE, 1982.

[5] Rabin, Michael O. "How To Exchange Secrets with Oblivious Transfer." IACR Cryptol. ePrint Arch. 2005.187 (2005).

[6] Rabin, Tal, and Michael Ben-Or. "Verifiable secret sharing and multiparty protocols with honest majority." Proceedings of the twenty-first annual ACM symposium on Theory of computing. 1989.
[7] Cramer, Ronald, Ivan Damgård, and Jesper B. Nielsen. "Multiparty computation from threshold homomorphic encryption." International conference on the theory and applications of cryptographic techniques. Springer, Berlin, Heidelberg, 2001.
[8] Rivest, Ronald L., Len Adleman, and Michael L. Dertouzos. "On data banks and privacy homomorphisms." Foundations of secure computation 4.11 (1978): 169-180.
[9] Dwork, Cynthia, et al. "Calibrating noise to sensitivity in private data analysis." Theory of cryptography conference. Springer, Berlin, Heidelberg, 2006.
[10] McSherry, Frank, and Kunal Talwar. "Mechanism design via differential privacy." 48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07). IEEE, 2007.
[11] Hard, Andrew, et al. "Federated learning for mobile keyboard prediction." arXiv preprint arXiv:1811.03604 (2018).
[12] Cheng, Kewei, et al. "Secureboost: A lossless federated learning framework." IEEE Intelligent Systems (2021).
[13] Bhagoji, Arjun Nitin, et al. "Analyzing federated learning through an adversarial lens." International Conference on Machine Learning. PMLR, 2019.
[14] Zhuo, Hankz Hankui, et al. "Federated reinforcement learning." (2019).
[15] Smith, Virginia, et al. "Federated multi-task learning." arXiv preprint arXiv:1705.10467 (2017).
[16] Yang, Qiang, et al. "Federated learning." *Synthesis Lectures on Artificial Intelligence and Machine Learning* 13.3 (2019): 1-207.
[17] Yang, Wensi, et al. "Ffd: A federated learning based method for credit card fraud detection." *International conference on big data*. Springer, Cham, 2019.
[18] Shingi, Geet. "A federated learning based approach for loan defaults prediction." *2020 International Conference on Data Mining Workshops (ICDMW)*. IEEE, 2020.
[19] Rieke, Nicola, et al. "The future of digital health with federated learning." *NPJ digital medicine* 3.1 (2020): 1-7.
[20] Liu, Yi, et al. "Privacy-preserving traffic flow prediction: A federated learning approach." *IEEE Internet of Things Journal* 7.8 (2020): 7751-7763.