WeBank 微众银行 | FedAI

联邦学习白皮书v2.0

微众银行人工智能部

电子商务与电子支付国家工程实验室

鹏城实验室

平安科技

腾讯研究院

云计算与大数据研究所

招商金融科技

·联合发布·

2020 年 4 月

电子商务与电子支付国家工程实验室 | 鹏城实验室 PengCheng Laboratory | 平安科技 PING AN TECHNOLOGY | 腾讯研究院 Tencent Research Institute | CAICT 中国信息通讯研究院 | CMGFintech 招商金融科技

目录 Contents

目录 Contents

第一章
联邦学习背景和重要性

1.1 人工智能发展现状

从 1955 年达特茅斯会议开始，人工智能经过两起两落的发展，迎来了第三个高峰期。第一个高峰期的出现是因为人们看到了 AI 的希望，也就是自动化算法对提高效率的希望，但是受算法能力的限制，机器不能完成大规模数据训练和复杂任务，AI 进入了第一个低谷。第二个高峰来自于霍普菲尔特神经网络的提出，以及 BP 算法实现了神经网络训练的突破，使得大规模神经网络训练成为可能。但是这时却发现算力和数据不够，专家系统的设计跟不上工业的成长需求，引发了 AI 的第二个低谷。2006 年，深度学习神经网络被提出，加上近年来算法和算力的巨大提升和大数据的出现，人工智能迎来了第三个高峰。2016 年的 AlphaGo，其总计使用了 30 万盘棋局作为训练数据并且接连战胜两位人类职业围棋选手，我们真正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自动驾驶、医疗、金融等更多、更复杂、更前沿的领域施展拳脚。

AlphaGo 的巨大成功使得人们自然而然地希望像这种大数据驱动的人工智能会在各行各业得以实现。但是真实的情况却让人非常失望：除了有限的几个行业，更多领域存在着数据有限且质量较差的问题，不足以支撑人工智能技术的实现。更多的应用领域有的只是小数据，或者质量很差的数据。这种“人工智能到处可用”的错误的认知会导致很严重的商业后果。一个案例是 IBM 的沃森，一个非常有名的问答（QA）系统，即给一个问题 Q，它能很精准找到答案 A。沃森可以用一个高维的表示来表达这个问题 Q，这种表示可以比喻为成物理学里的光谱，棱镜把一束光分解成不同频率的光，形成光谱。有了这个光谱以后，可以和答案库里对应答案，概率相应高的就是可能的答案。整个流程应该说非常简单，但问题就是要有一个很健全的答案库。IBM 在电视大赛上取得了成功之后，就把这个应用在一些听起来比较好的垂直领域——医疗领域。然而，最近在一个美国的癌症治疗中心，发现这个应用非常不理想，从而导致了这个项目的失败。我们可以看一看在医疗领域，这些领域里的问题和答案来自哪里?比如输入有病症、基因序列、病理报告、各种各样的检测、各种论文，沃森的任务是利用这些数据来做诊断，帮助医生。但是，经过一段时间的实践发现，这些数据的来源远远不够，导致了系统效果很差。医疗领域需要非常多的标注数据，而医生的时间却非常宝贵，不能像其他的一些计算机视觉应用一样，可以由大众普通人来完成数据标注。所以在医疗这样的专业领域，这种标注的数据非常有限。有人估计，把医疗数据放在第三方公司标注，需要动用 1 万人用长达 10 年的时间才能收集到有效的数据。这就说明，在这些领域，即使动用很多人来做标注，数据也不够。这就是我们面临的现实。

同时数据源之间存在着难以打破的壁垒，一般情况下人工智能的所需要的数据会涉及多个领域，例如在基于人工智能的产品推荐服务中，产品销售方拥有产品的数据、用户购买商品的数据，但是没有用户购买能力和支付习惯的数据。在大多数行业中，数据是以孤岛的形式存在的，由于行业竞争、隐私安全、行政手续复杂等问题，即使是在同一个公司的不同部门之间实现数据整合也面临着重重阻力，在现实中想要将分散在各地、各个机构的数据进行整合几乎是不可能的，或者说所需的成本是巨大的。

1.2 人工智能面临的挑战

另一方面，随着大数据的进一步发展，重视数据隐私和安全已经成为了世界性的趋势。每一次公众数据的泄露都会引起媒体和公众的极大关注，例如最近 Facebook 的数据泄露事件就引起了大范围的抗议行动。同时各国都在加强对数据安全和隐私的保护，欧盟 2018 年正式施行的法案《通用数据保护条例》（General Data Protection Regulation, GDPR）表明，对用户数据隐私和安全管理的日趋严格将是世界趋势。这给人工智能领域带来了前所未有的挑战，研究界和企业界目前的情况是收集数据的一方通常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 GDPR，并可能遭到法案严厉的惩罚。同样，中国在 2017 年起实施的《中华人民共和国网络安全法》[1] 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。这些法规的建立在不同程度上对人工智能传统的数据处理模式提出了新的挑战。在这个问题上，人工智能的学界和企业界，目前并无较好的解决方案来应对这些挑战。

1.3 数据隐私保护可行性解决方案

要解决大数据的困境，仅仅靠传统的方法已经出现瓶颈。两个公司简单的交换数据在很多法规包括 GDPR 是不允许的。用户是原始数据的拥有者，在用户没有批准的情况下，公司间不能交换数据。其次，数据建模使用的目的，在用户认可前不可以改变。所以，过去的许多数据交换的尝试，例如数据交易所的数据交换，也需要巨大的改变才能合规。同时，商业公司所拥有的数据往往有巨大的潜在价值。两个公司甚至公司间的部门都要考虑

利益的交换，在这个前提下，往往这些部门不会把数据与其他部门做简单的聚合。这将导致即使在同一个公司内，数据也往往以孤岛形式出现。

如何在满足数据隐私、安全和监管要求的前提下，设计一个机器学习框架，让人工智能系统能够更加高效、准确地共同使用各自的数据，是当前人工智能发展的一个重要课题。我们倡议把研究的重点转移到如何解决数据孤岛的问题。我们提出一个满足隐私保护和数据安全的一个可行的解决方案，叫做联邦学习 [2-4]。

联邦学习是：

- 各方数据都保留在本地，不泄露隐私也不违反法规；
- 多个参与者联合数据建立虚拟的共有模型，并且共同获益的体系；
- 在联邦学习的体系下，各个参与者的身份和地位平等；
- 联邦学习的建模效果和将整个数据集放在一处建模的效果相同，或相差不大（在各个数据的用户对齐（user alignment）或特征（feature alignment）对齐的条件下）；
- 迁移学习是在用户或特征不对齐的情况下，也可以在数据间通过交换加密参数达到知识迁移的效果。

联邦学习使多个参与方在保护数据隐私、满足合法合规要求的前提下继续进行机器学习，解决数据孤岛问题。

第二章
联邦学习定义及价值分析

2.1 联邦学习概述

什么是联邦学习呢?举例来说，假设有两个不同的企业 A 和 B，它们拥有不同的数据。比如，企业 A 有用户特征数据；企业 B 有产品特征数据和标注数据。这两个企业按照上述 GDPR 准则是不能粗暴地把双方数据加以合并的，因为数据的原始提供者，即他们各自的用户并没有机会来同意这样做。假设双方各自建立一个任务模型，每个任务可以是分类或预测，而这些任务也已经在获得数据时有各自用户的认可。那现在的问题是如何在 A 和 B 各端建立高质量的模型。但是，由于数据不完整（例如企业 A 缺少标签数据，企业 B 缺少特征数据），或者数据不充分（数据量不足以建立好的模型），那么，在各端的模型有可能无法建立或效果并不理想。联邦学习是要解决这个问题：它希望做到各个企业的自有数据不出本地，而联邦系统可以通过加密机制下的参数交换方式，即在不违反数据隐私法规情况下，建立一个虚拟的共有模型。这个虚拟模型就好像大家把数据聚合在一起建立的最优模型一样。但是在建立虚拟模型的时候，数据本身不移动，也不泄露隐私和影响数据合规。这样，建好的模型在各自的区域仅为本地的目标服务。在这样一个联邦机制下，各个参与者的身份和地位相同，而联邦系统帮助大家建立了“共同富裕”的策略。这就是为什么这个体系叫做“联邦学习”。

上述实例阐述了联邦学习的基本思想，下文将规范联邦学习的定义，介绍联邦学习的公共价值和商业价值，并阐明联邦学习与现有研究的关系。

2.2 联邦学习的定义

为了进一步准确地阐述联邦学习的思想，我们将其定义如下：

在进行机器学习的过程中，各参与方可借助其他方数据进行联合建模。各方无需共享数据资源，即数据不出本地的情况下，进行数据联合训练，建立共享的机器学习模型。

联邦学习系统的约束条件为：

$$|V\_FED-V\_SUM|<\delta$$

式中：

V_FED —— 联邦学习模型的效果；

V_SUM —— 传统方法（数据聚合方法）所建模型的效果；

δ —— 有界正数。

图 1 联邦学习架构

2.3 联邦学习的公共价值

毫无疑问，如今我们正经历互联网第四次信息革命，坐拥海量的信息与数据。这些数据如果能够用 AI 的方式进行解读，将会为人类日常生活带来颠覆性变革。

联邦学习作为未来 AI 发展的底层技术，它依靠安全可信的数据保护措施下连接数据孤岛的模式，将不断推动全球 AI 技术的创新与飞跃。随着联邦学习在更大范围和更多行业场景的渗透及应用，它在更高层面上对各类人群、组织、行业和社会都将产生巨大影响，联邦学习的公共价值主要体现在以下几个方面：

图 2 联邦学习公共价值

加速人工智能技术创新发展

人工智能技术当前已形成汇聚了全球技术、资金、人才和影响力等多元资源的产业生态，而作为 AI 建模底层不可或缺的核心技术，联邦学习将真正助力大数据实现价值，在数据不出本地的环境下带动 AI 各领域在各行业的深度融合，使得人工智能技术能够扫清数据障碍，不断迭代成长和创新。

保障隐私信息及数据安全

联邦学习可做到个体的自有数据不出本地，联邦系统通过加密机制下的参数交换方式，在不违反数据隐私保护法规的情况下，建立一个虚拟的共有模型。建立虚拟模型时，数据本身不移动，也不会泄露用户隐私或影响数据规范，充分保障了个体隐私信息及数据安全。

促进全社会智能化水平提升

基于联邦学习的 AI 技术将更安全地融入社会基础设施和生活中，它不仅能辅助人类的工作及生活，也逐步改变人类的认知模式，从而推动社会经济及发展。

2.4 联邦学习的商业价值

联邦学习技术是一种“合作共赢”的模式，对商业利益而言极具价值。在这样一个联邦机制下，各个参与者的身份和地位相同，而联邦系统帮助大家建立了“共同富裕”的策略。这就是为什么这个体系叫做“联邦学习”。从商业角度而言，联邦学习的主要价值有：

图 3 联邦学习商业价值

带动跨领域的企业级数据合作，智能策略辅助市场布局及竞争力提升

联邦学习作为 AI 发展的底层技术，能够帮助到企业参与到新的全球化、泛行业化的协作网络和联邦生态中，通过跨领域的企业界数据合作，更有效地训练模型辅助自身市场布局、策略优化，从而提升竞争力。联邦学习能在技术层面帮助企业更好地确立自身合作与竞争策略，以此形成联邦中的独有生态，从而更好推动企业良性发展。

催生基于联合建模的新业态和模式

通过联邦生态及联邦学习在其他领域的应用拓展，将不断影响和改变合作中提供方、需求方的关系，重定义各方合作者的身份、服务方式和盈利方式，催生出基于联合建模的全新业态及模式。

降低技术提升成本和促进创新技术发展

联邦学习技术成体系可复用的解决方案能够有效降低技术应用的门槛，扩大技术应用的范围和广度，这使得广大泛 AI 行业企业机构能够为不同客户提供更加丰富的产品及服务，同时去除数据安全隐忧的 AI 大环境将有助于创新型技术的进一步飞跃，在提升效率和获得成长的同时，实现自身发展。

2.5 联邦学习与现有研究的关系

作为一种全新的技术，联邦学习在借鉴一些成熟技术的同时也具备了一定的独创性。下面我们就从多个角度来阐释联邦学习和其他相关概念之间的关系。

联邦学习与差分隐私理论的区别

联邦学习的特点使其可以被用来保护用户数据的隐私，但是它和大数据、数据挖掘领域中常用的隐私保护理论如差分隐私保护理论（Differential Privacy）[5]、k 匿名（k-Anonymity）[6] 和 l 多样化（l-Diversity）[7] 等方法还是有较大的差别的。首先联邦学习与传统隐私保护方法的原理不同，联邦学习通过加密机制下的参数交换方式保护用户数据隐私，加密手段包括同态加密 [8-12] 等。与 Differential Privacy 不同，其数据和模型本身不会进行传输，因此在数据层面上不存在泄露的可能，也不违反更严格的数据保护法案如 GDPR 等。而差分隐私理论、

k 匿名和 l 多样化等方法是通过在数据里加噪音，或者采用概括化的方法模糊某些敏感属性，直到第三方不能区分个体为止，从而以较高的概率使数据无法被还原，以此来保护用户隐私。但是，从本质上来说这些方法还是进行了原始数据的传输，存在着潜在被攻击的可能性，并且在 GDPR 等更严格的数据保护法案下这种数据隐私的保护方式可能不再适用。与之对应的，联邦学习是对用户数据隐私保护更为有力的手段。

联邦学习与分布式机器学习的区别

横向联邦学习中多方联合训练的方式与分布式机器学习（Distributed Machine Learning）有部分相似的地方。分布式机器学习涵盖了多个方面，包括把机器学习中的训练数据分布式存储、计算任务分布式运行、模型结果分布式发布等，参数服务器（Parameter Server）[13] 是分布式机器学习中一个典型的例子。参数服务器作为加速机器学习模型训练过程的一种工具，它将数据存储在分布式的工作节点上，通过一个中心式的调度节点调配数据分布和分配计算资源，以便更高效的获得最终的训练模型。而对于联邦学习而言，首先在于横向联邦学习中的工作节点代表的是模型训练的数据拥有方，其对本地的数据具有完全的自治权限，可以自主决定何时加入联邦学习进行建模，相对地在参数服务器中，中心节点始终占据着主导地位，因此联邦学习面对的是一个更复杂的学习环境；其次，联邦学习则强调模型训练过程中对数据拥有方的数据隐私保护，是一种应对数据隐私保护的有效措施，能够更好地应对未来愈加严格的数据隐私和数据安全监管环境。

联邦学习与联邦数据库的关系

联邦数据库系统（Federated Database System）[14] 是将多个不同的单元数据库进行集成，并对集成后的整体进行管理的系统。它的提出是为了实现对多个独立的数据库进行相互操作。联邦数据库系统对单元数据库往往采用分布式存储的方式，并且在实际中各个单元数据库中的数据是异构的，因此，它和联邦学习在数据的类型与存储方式上有很多相似之处。但是，联邦数据库系统在各个单元数据库交互的过程中不涉及任何隐私保护机制，所有单元数据库对管理系统都是完全可见的。此外，联邦数据库系统的工作重心在包括插入、删除、查找、合并等各种数据库基本操作上面，而联邦学习的目的是在保护数据隐私的前提下对各个数据建立一个联合模型，使数据中蕴含的各种模式与规律更好地为我们服务。

联邦学习与区块链的关系

区块链是一个基于密码学安全的分布式账本，其方便验证，不可篡改。区块链 2.0 是一个去中心化的应用，通过使用开源的代码及分布式的存储和运行，保证极高的透明度和安全性，使数据不会被篡改。区块链的典型应用包括比特币（BTC）、以太坊（ETH）等。区块链与联邦学习都是一种去中心化的网络，区块链是一种完全 P2P（peer to peer）的网络结构，在联邦学习中，第三方会承担汇聚模型、管理等功能。联邦学习与区块链中，均涉及到密码学、加密算法等基础技术。根据技术的不同，区块链技术使用的加密算法包括哈希算法，非对称加密等; 联邦学习中使用同态加密等。从数据角度上看，区块链上通过加密的方式在各个节点上记录了完整的数据，而联邦学习中，各方的数据均仅保留在本地。从奖励机制上看，区块链中，不同节点之间通过竞争记账来获得奖励; 在联邦学习中，多个参与方通过共同学习，提高模型训练结果，依据每一方的贡献来分配奖励。

联邦学习与多方安全计算的关系

在联邦学习中，用户的隐私与安全是重中之重。为了保护用户隐私，防止联邦学习应用被恶意方攻击，多方安全计算技术可以在联邦学习中被应用，成为联邦学习技术框架中的一部分。学术界已经展开利用多方安全计算来增强联邦学习的安全性的研究。McMahan [15] 指出，联邦学习可以通过差分隐私，多方安全计算，或它们的结合等技术来提供更强的安全保障。Bonawitz [16] 指出，联邦学习中，可以利用多方安全计算以安全的方式计算来自用户设备的模型参数更新的总和。Truex[17] 中提出了一种利用差分隐私和多方安全计算来保护隐私的联邦学习方法。Liu[18] 提出将加性同态加密 (AHE) 应用于神经网络的多方计算。微众银行提出的开源联邦学习框架 FATE [19] 中包含了多方安全计算的相关算子，方便应用方对多方安全计算进行高效的开发。

第三章
联邦学习分类

联邦学习分类

上述对联邦学习的定义并没有讨论如何具体地设计一种联邦学习的实施方案。在实际中，孤岛数据具有不同分布特点，根据这些特点，我们可以提出相对应的联邦学习方案。下面，我们将以孤岛数据的分布特点为依据对联邦学习进行分类。

考虑有多个数据拥有方，每个数据拥有方各自所持有的数据集 D_i 可以用一个矩阵来表示。矩阵的每一行代表一个用户，每一列代表一种用户特征。同时，某些数据集可能还包含标签数据。如果要对用户行为建立预测模型，就必须要有标签数据。我们可以把用户特征叫做 X，把标签特征叫做 Y。比如，在金融领域，用户的信用是需要被预测的标签 Y；在营销领域，标签是用户的购买愿望 Y；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，往往会遇到这样的情况：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，以包含两个数据拥有方的联邦学习为例，数据分布可以分为以下三种情况：

- 两个数据集的用户特征(X1,X2,...)重叠部分较大，而用户(U1, U2...)重叠部分较小；
- 两个数据集的用户(U1, U2...)重叠部分较大，而用户特征(X1,X2,...)重叠部分较小；
- 两个数据集的用户(U1, U2...)与用户特征重叠(X1,X2,...)部分都比较小。

为了应对以上三种数据分布情况，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，如图 4 所示。

图 4 联邦学习的分类

3.1 横向联邦学习

在两个数据集的用户特征重叠较多而用户重叠较少的情况下，我们把数据集按照横向（即用户维度）切分，并取出双方用户特征相同而用户不完全相同的那部分数据进行训练。这种方法叫做横向联邦学习。比如有两家不同地区银行，它们的用户群体分别来自各自所在的地区，相互的交集很小。但是，它们的业务很相似，因此，记录的用户特征是相同的。此时，就可以使用横向联邦学习来构建联合模型。Google 在 2017 年提出了一个针对安卓手机模型更新的数据联合建模方案 [15,20]：在单个用户使用安卓手机时，不断在本地更新模型参数并将参数上传到安卓云上，从而使特征维度相同的各数据拥有方建立联合模型的一种联邦学习方案。

3.2 纵向联邦学习

在两个数据集的用户重叠较多而用户特征重叠较少的情况下，我们把数据集按照纵向（即特征维度）切分，并取出双方用户相同而用户特征不完全相同的那部分数据进行训练。这种方法叫做纵向联邦学习。比如有两个不同机构，一家是某地的银行，另一家是同一个地方的电商。它们的用户群体很有可能包含该地的大部分居民，因此用户的交集较大。但是，由于银行记录的都是用户的收支行为与信用评级，而电商则保有用户的浏览与购买历史，因此它们的用户特征交集较小。纵向联邦学习就是将这些不同特征在加密的状态下加以聚合，以增强模型能力的联邦学习。目前，逻辑回归模型，树型结构模型和神经网络模型等众多机器学习模型已经逐渐被证实能够建立在这个联邦体系上。

3.3 联邦迁移学习

在两个数据集的用户与用户特征重叠都较少的情况下，我们不对数据进行切分，而可以利用迁移学习 [21] 来克服数据或标签不足的情况，这种方法叫作联邦迁移学习 [22]。

比如有两个不同机构，一家是位于中国的银行，另一家是位于美国的电商。由于受到地域限制，这两家机构的用户群体交集很小。同时，由于机构类型的不同，二者的数据特征也只有小部分重合。在这种情况下，要想进行有效的联邦学习，就必须引入迁移学习，来解决单边数据规模小和标签样本少的问题，从而提升模型的效果。

第四章
联邦学习框架

4.1 联邦学习开源框架介绍

目前业界中主要的联邦学习框架有 FATE[23]，TensorFlow Federated[24]，PaddleFL[25]，Pysyft[26] 等。

2019 年 2 月，微众银行开源 FATE 开源项目，截止 2019 年 12 月发布 FATE v1.2 版本，覆盖横向联邦学习，纵向联邦学习，联邦迁移学习，得到了社区内广泛的关注与应用。同时，FATE 提供 20 多个联邦学习算法组件，涵盖 LR，GBDT, DNN 等主流算法，覆盖常规商业应用场景建模需求。尤其值得注意的是，FATE 提供了一站式联邦模型服务解决方案，涵盖联邦特征工程，联邦机器学习模型训练，联邦模型评估，联邦在线推理，相比其他开源框架，在工业应用上有显著的优势。

OpenMinded 开源的 Pysyft 框架，较好地支持横向联邦学习。该框架同时支持 Tensorflow，Keras，Pytorch，为使用人员快速上手提供了较多的选择。Pysyft 提供了安全加密算子，数值运算算子，及联邦学习算法，用户也可以高效搭建自己的联邦学习算法。相比较 FATE，OpenMinded 尚未提供高效的部署方案及 serving 端解决方案，相比工业应用，更适合作为高效的学术研究、原型开发的工具。

谷歌开源的 TensorFlow Federated 框架，截止 2019 年 12 月已发布至 0.11 版本，较好地支持横向联邦学习。其中，可以通过 Federated Learning(FL) API，与 Tensorflow/Keras 交互，完成分类、回归等任务。用户也可以通过其提供的 Federated Core (FC) API，通过在强类型函数编程环境中将 TensorFlow 与分布式通信运算符相结合，简洁地表达新的联合算法。目前 TensorFlow Federated 在安全加密算子上缺少开放实现，同时缺少对线上生产的完善支撑。

2019 年 11 月，百度宣布开源其联邦学习框架 PaddleFL。PaddleFL 开源框架中包含了 DiffieHellman 等安全算子，及 LR 等机器学习算法。由于其开源时间较短，算子丰富程度逊于上述三个框架。PaddleFL 的优势在于通过与百度机器学习开源框架 PaddlePaddle 的交互，吸引相关生态开发者加入开发。

联邦学习开源框架对比如下：

开源框架	FATE	TensorFlow Federated	PaddleFL	Pysyft
受众定位	工业产品/学术研究	学术研究	学术研究	学术研究
牵头公司/机构	微众银行	Google	百度	OpenMined
联邦学习类型	横向联邦学习 纵向联邦学习 联邦迁移学习	横向联邦学习	横向联邦学习 纵向联邦学习	横向联邦学习
联邦特征工程算法	特征分箱 特征选择 特征相关性分析 支持	不支持	不支持	不支持
机器学习算法	LR，GBDT，DNN等	LR，DNN等	LR，DNN等	LR，DNN等
安全协议	同态加密，SecretShare，RSA，DiffieHellman	DP	DP	同态加密，SecretShare
联邦在线推理	支持	不支持	不支持	不支持
Kubernetes	支持	不支持	不支持	不支持
代码托管平台	Github(https://github.com/FederatedAI/FATE)	Github(https://github.com/tensorflow/federated)	Github(https://github.com/PaddlePaddle/PaddleFL)	Github(https://github.com/OpenMined/PySyft)

4.2 联邦学习企业级架构——FATE

2019 年 2 月，微众银行 AI 团队对外发布自主研发的开源项目 FATE（Federated AI Technology Enabler），作为全球首个联邦学习开源框架，FATE 为联邦 AI 生态提供了一种安全计算框架。

FATE 提供了一种基于数据隐私保护的分布式安全计算框架，为机器学习、深度学习、迁移学习算法提供高性能的安全计算支持，支持同态加密、SecretShare、DiffieHellman 等多种多方安全计算协议。同时，FATE 提供了一套友好的跨域交互信息管理方案，解决了联邦学习信息安全审计难的问题。简单易用的开源工具平台能有效帮助多个机构在满足用户隐私保护、数据安全和政府法规的前提下，进行多方数据合作。

目前 FATE 已在信贷风控、客户权益定价、监管科技等领域推动应用落地。

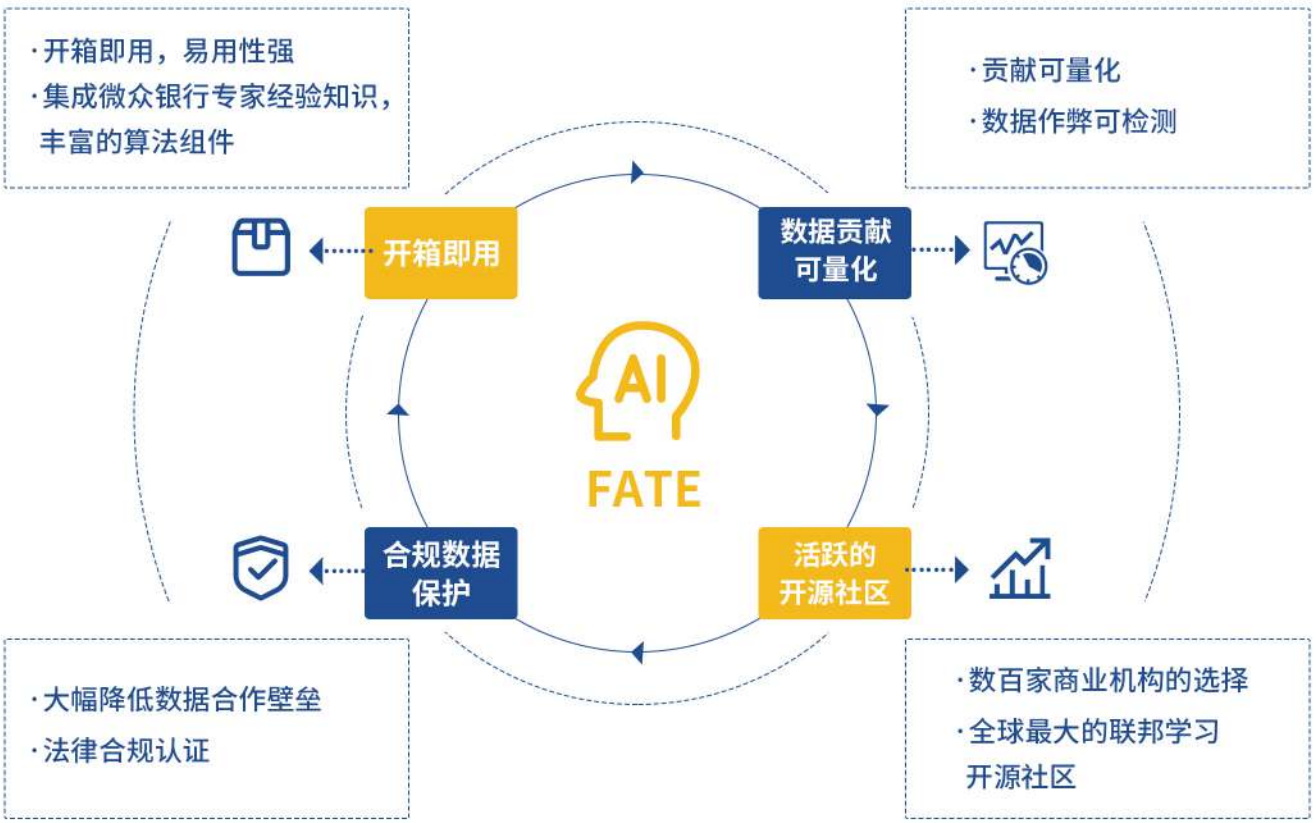

图 5 FATE 框架优势

图 6 FATE 2019 里程碑

FATE技术框架

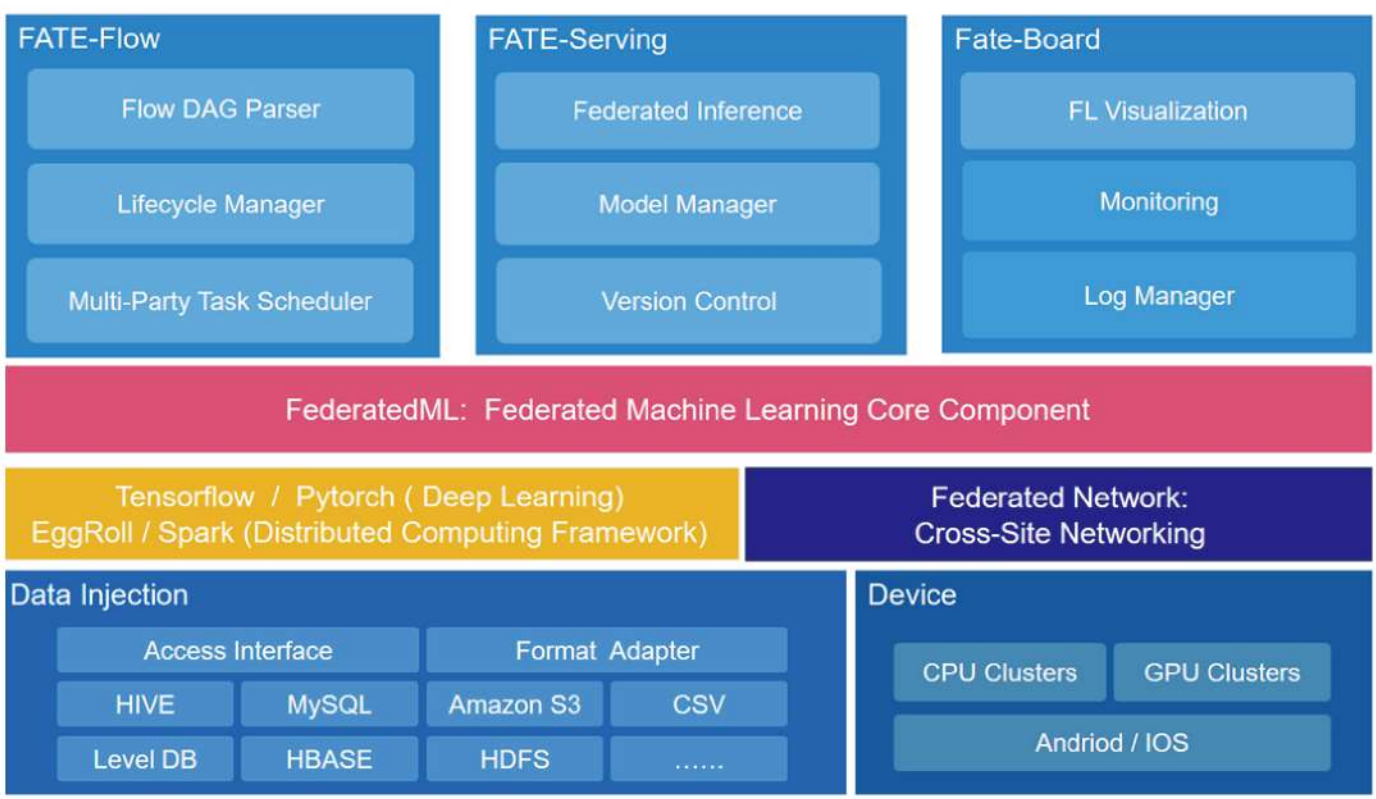

图 7 FATE 技术框架

FederatedML：

联邦学习算法功能组件，包括了许多常见机器学习算法联邦化实现。所有模块均采用模块化的解耦的方式进行开发，从而增强可扩展性。

主要功能：

- 联邦样本对齐：纵向样本ID对齐，包括基于RSA+哈希等对齐方式
- 联邦特征工程：联邦采样，联邦特征分箱，联邦特征选择，联邦相关性，联邦统计等
- 联邦机器学习：联邦LogisticRegression，LinearRegression，PossionRegression，联邦SecureBoost,联邦DNN，联邦迁移学习等
- 多方安全计算协议：提供多种安全协议，包括同态加密，SecretShare，RSA，DiffieHellman等

FATE-Flow:

联邦学习建模Pipeline 调度和生命周期管理工具，为用户构建端到端的联邦学习pipeline生产服务。

主要功能：

- 联邦建模Pipeline DAG Parser
- 联邦建模任务生命周期管理
- 联邦建模任务多方协同调度
- 联邦多方模型管理、模型版本管理
- 联邦建模过程数据、指标、模型等输入输出实时跟踪

FATE-Board：

联邦学习建模的可视化工具，为终端用户可视化和度量模型训练的全过程。支持对模型训练过程全流程的跟踪、统计和监控等，并为模型运行状态、模型输出、日志追踪等提供了丰富的可视化呈现，帮助用户简单而高效地深入探索模型与理解模型。

主要功能：

- 联邦建模任务生命周期过程可视化
- 联邦模型可视化
- 评估报告可视化

FATE-Serving：

高性能可扩展的联邦学习在线模型服务。

主要功能：

- 高性能在线联邦模型推理算法
- 在线联邦模型管理
- 联邦学习在线推理pipeline

KubeFATE：

通过把FATE的所有组件用容器的形式封装，实现了使用Docker Compose或Kubernetes（Helm Charts）来部署。现代应用以DevOps方式开发，基于容器部署应用的优势相当明显，应用不仅可以无差别地运行在支持容器的平台上，还可以按需灵活地实现多实例水平扩展。通过KubeFATE项目，开发者可以轻松地在公有云或私有云中部署FATE项目。

第五章
联邦学习应用实例

联邦学习是否有价值，取决于联邦学习的关键应用场景。只有通过联邦学习的应用实例化，才能发现联邦学习在发展中所遇到的各种挑战和机遇。下文我们将介绍七个联邦学习的典型应用场景。

5.1 联邦车险定价

背景及需求

我国保险行业已进入快速发展阶段。但日益激烈的市场竞争环境、快速发展的新产品种类，对保险业的风险识别、精准定价能力提出了更高的要求，传统的保险风险识别不能完全满足业务需求，甚至影响产品盈利能力。

以车险为例，传统车险的定价方式是从车定价，根据车的品质来确认保费，好车保费远高于普通车，但实际上，车辆使用情况、行车区域环境等方面的内容也是影响保期内赔付风险的重要因素，对车险保费定价有着至关重要的影响。因此，从车定价到从人定价，是消费者市场的大趋势。但对强监管的保险行业而言，影响定价精确性的数据分布分散，只有垂直场景的交易数据，新客数据表现也较少，缺乏有效机制将数据链接整合起来，难以实现精准定价。

此外，随时代发展，保险行业产品同质化严重，营销方面无法进一步精准触达，也影响着保险企业及机构的未来发展。

基于联邦学习的解决方案

车险定价方面，针对从人、从车、从行为等数据分散在不同的公司，数据无法出库，无法直接进行聚合并建模的症结，引入联邦学习机制建模，在保护各合作机构企业用户隐私数据不出库的前提下，安全合规接入多方数据源，打破数据壁垒。

其中，面对产品同质化严重导致定价僵硬问题，着重引入多维度多标签的互联网行为数据，完善用户画像，提供个性化定价服务，推动智能营销决策。而针对新用户道德风险把控等场景，引入安全大数据，有效识别恶意用户和欺诈骗保行为。

联邦学习解决方案效益

基于联邦学习建立的数据模型具有丰富的风险特征体系，能有效判别风险、预测赔付成本，并提供个性化定价服务。行业定价准确率大幅提升，总比例超 90%。

此外，方案通过符合隐私保护条例的大数据模型，进一步升级智能营销服务体系，提升更加精准化、定制化的综合金融服务能力，精准挖掘高潜客户，从而协助保险机构及企业解决新客覆盖问题。

行业前景

数据对保险企业及机构来说，可以更好地读取更多关于保险用户市场的信息。基于联邦学习的智慧金融能实现保护数据安全基础上，打破数据壁垒，实现保险行业上下游业务场景深度结合，解决保险企业、中介机构及代理人数据孤立、客户体验差等困扰，通过降低营销运营成本，提升数据服务效率，提高其全线质量，助力保险市场数字化健康有序发展。

5.2 联邦信贷风控

背景及需求

在信贷风控领域中，因信审过程中需要调用不同数据接口，故面临着单客户的信贷审核成本较为高昂的情况，例如对于消费金融、小微企业信贷等行业，审核中的核身、征信等接口的调用费用极高。

此外，银行等金融机构在面对小微企业信贷需求时，缺乏企业经营状况等有效数据，而导致小微企业融资难、融资贵、融资慢。同样消费金融类企业机构在面对风控时，缺乏互联网行为画像等有效数据，随着贷款客户逐步下沉，客户的信用资质也更加参差不齐。

如何结合 AI 赋能传统金融行业，合法合规地从多源数据中综合判断贷款客户的资质和信用情况，如小微企业的盈利能力、收入波动和成长潜力等，以及用户的消费行为、消费能力和兴趣置好等。怎样让小微企业共享普惠金融，怎样拉近消费金融公司到支付和消费场景的距离，是亟待解决的重大问题。

基于联邦学习的解决方案

通过联邦数据网络进行信贷风控增强，助力消费金融公司信贷预审。从风险源头切入，帮助信贷公司过滤信贷黑名单或明显没有转化的贷款客户，进一步降低贷款审批流程后期的信审成本。

针对消费金融企业机构 Y 样本量不足、好坏样本区分度不够、样本分布偏离正态分布等问题，解决方案设立长时累积数据，并协助从相关合作方——信贷机构获取更多同类 Y 数据，增加样本量，此外，在样本采集过程中也持续进行反复筛选及重采，与数据双方实时同步数据资料，剔除 18 年 X 对应 19 年 Y 等数据参差情况，使全量数据时刻保持基本符合正态分布。具体实施方面，解决方案先行使用联邦学习云服务进行业务冷启动，并通过建立业务及 AI 模型闭环，小样本建模，后期持续迭代优化模型的方式，实现项目数字化，便于消费金融业务方及信贷合作方能够持续积累业务数据优化联邦模型。

针对小微企业信贷评审数据稀缺、不全面、历史信息沉淀不足等问题，解决方案为设立多源数据融合机制，包括小微企业的交易数据、税务数据、工商信息等，并协助相关金融机构获取更多维度的 X 数据，丰富特征体系。此外，在特征体系构建的过程中，确保数据提供方的数据安全及隐私保密等问题，并共同提升模型的有效性。具体实施方面，解决方案先行使用联邦学习云服务进行样本对齐，并通过纵向联邦学习建模，以及参数输出，便于合作方能够观察到自身数据对于优化联邦模型的有效性。

联邦学习解决方案效益

通过合法合规的多维度联邦数据建模，风控模型效果约可提升 12%，消费金融类企业机构有效节约了信贷审核成本，整体成本预计下降 5%-10%，并因数据样本量的提升和丰富，风控能力进一步增强。而对合作方信贷机构而言，信贷风控能力也大幅度提升，通过初审筛选掉黑名单和不可能转化贷款客户，在“信审漏斗第一步”减去无效客户，从而在信贷预审阶段使单接口调用成本预计节省 20-30%，有效控制了信贷审核成本。

通过联邦数据建模，金融机构以及信贷机构的数据孤岛得以连接，合法合规地最大化了自有数据价值，金融机构进一步靠近了支付及消费场景，信贷机构核心竞争力也获得增强。

行业前景

基于联邦学习的智能风控有利于促进基于数据安全联合建模的 AI 技术应用落地，更好地支撑消费金融行业的价值创造，并提升此类行业的风险控制能力。同时可以通过联邦学习提升金融科技公司对金融机构各项业务的服务能力。

目前，联邦学习已实现可参与至风控流程各环节，包括反欺诈、白名单初筛、信贷预审、贷中和贷后预警评分等。根据业务企业及机构需求，可进行多维度合作，后续，联邦学习将通过深入信贷风控审核主流程，进一步用联邦建模渗透到信审各环节，实现数据隐私保护下的数据连接及合作。

5.3 联邦销量预测

背景及需求

随着我国经济发展，消费升级进程也在持续加快，居民消费结构经历不断优化，服务消费需求快速增长，商品零售增速也随之持续放缓。据国家统计局数据，2018 年我国社会消费品零售总额为 38.09 万亿元，环比增长 4.02%，环比增速比上年下降 6.19 个百分点，环比增速已连续 5 年下降。

面对实体零售业存量巨大、增速放缓的现状，新零售应运而生，如何阻断行业用工成本逐年上升侵蚀毛利的趋势?怎样解决生鲜业务占比逐年攀升，但损耗却始终居高不下的问题?基于合法合规前提下的数字化经营成为零售企业应对变化和挑战的重要手段。

基于联邦学习的解决方案

以生鲜损耗为例，解决方案通过动态盘点、排面管理及时出清、爆品预测及推荐等，可进一步实现门店生鲜智能管控，完成货物新鲜度及折扣促销策略实时对应调整，推动止损策略及时反馈。同时，由大数据进行爆品提前预测，为零售企业提供有效爆品信息，便于企业联动进行营销策略优化，提升销售转化。

在这一智慧零售业务场景中，涉及到的数据特征主要包含用户购买能力，用户个人偏好，以及产品特点三部分，但是在实际应用中，这三种数据特征很可能分散在三个不同的部门或企业。例如，银行拥有用户购买能力的特征，社交网站拥有用户个人偏好特征，而购物网站则拥有产品特点的特征。这种情况下，我们面临两大难题：首先，出于保护用户隐私以及企业数据安全等原因，银行、社交网站和购物网站三方之间的数据壁垒是很难被打破的，智慧零售的业务部门无法直接把数据进行聚合并建模；其次，这三方的用户特征数据通常是异构的，传统的机器学习模型无法直接在异构数据上进行学习。

但利用联邦学习，不用导出企业数据，三方即可联合构建机器学习模型，既充分保护了用户隐私和数据安全，又能为用户提供个性化商品配货及营销策略，针对性的产品服务，实现多方共同受益。

联邦学习解决方案效益

通过科技赋能生鲜管理，帮助企业实现生鲜精细化运营，销量预测准确度达到 85% 以上，工程优化后，准确率达到 95% 以上。零售企业进一步降低损耗，并较大幅度提升了销量。此外，通过劳动力数字化，企业可实行跨岗灵活用工，高效利用场内外员工闲暇时间，充分调动员工工作积极性，促进高效生产，智能用工数字化平

台经企业落地，人均效率提升约27%。

经联邦建模数字化，零售类企业可合法合规地进行个性化产品服务，售卖模式得到进一步拓展。在服务好用户的同时，为精准营销打下了基础。用户可以体验到更贴合自身需求的购物模式，同时零售企业人力成本也得到控制，人效与坪效均获提升。

行业前景

基于联邦学习的智能用工系统能够提供一个覆盖劳动前、中、后全流程的完整的劳动力数字化平台，解决用工需求方及劳动力提供方之间的信息不对称，结合劳动力密集型行业的多样化场景及 AI、银行金融能力，为双方提供灵活性强、自由度高的无缝对接平台及银行级的清结算、信用累积等增值服务。

同时，还可为企业提供计划、采购、仓储、生产制造、销售 E2E 的 AI 数字化能力，从根本上颠覆线下企业的运营模式，帮助企业实现真正以用户需求为导向的智能化运营体系。

后续，联邦学习的这一应用落地模式可进一步拓展行业多样化，如制造业、仓储物流业、进出口贸易等垂直领域，从而将 AI、大数据、以及工程产品能力等，在数据隐私安全保护的规范性下，应用至国民经济各领域。

5.4 联邦视觉安防

背景及需求

眼下，中国城市的发展遵循其内在规律，已经走过了 40 年的“由农村到城市化”的演进，而接下来的几十年，城市的“智慧化”需求正逐渐成为城市发展新的动力。

在国内，有 83% 的地级城市、总计超过 500 个城市明确提出或正在建设智慧城市。作为智慧城市的重要组成部分，智慧安防是一大方向，传统安防常以摄像头收集基础数据，IT 系统以及多处理器处理信息，同时，设置监控室作为监控环节，再辅以人工进行检测非安全举动。这种非安全的排查流程较冗长，人力成本高且社区管理效率低。

此外，社区人员的出行情况存在不可预知性，特殊人群（如老人、吸毒、犯罪人员）的异常情况无法及时发现，社区安全应对滞后，且现有的异常定义依赖于人工规则，预警存在误差及漏判。虽然社区通过摄像头、门禁卡等

方式，收集了大量用户社区通行数据，但彼此之间互不关联，无数社区陷入支离破碎的信息孤岛，数据价值无法得到挖掘。

基于联邦学习的解决方案

智慧城市的发展离不开安防这一建设切入点，智慧安防是城市建设的第一步和核心模块。

预设算法训练模型进行事前预警，实时高精度重点拍摄、判断位置、识别动作以及分析行为，预测社区用户出行轨迹及用户异常出行，从而提升社区安全及社区管理效率。通过引入联邦学习，联合多社区数据进行安防模型建立，实现数个社区间信息的互联互通，让智能安防建立在同纬度的一张网上。

基于云计算及大数据分析，智慧安防还可以做到事后总结和自我学习，就像一个永远不知疲倦、不会退休的警官，不停的积累安防经验，持续提高事前预警能力。

联邦学习解决方案效益

在日常的场景应用中，大量来自于视频、传感器、信息软件的数据通过汇集整理分析，为社区提供更加安全、精确、更具有成长性的服务，从而帮助社区用户进行风险预测。

考虑到数据安全性，暂无法实现数据集中建模。但通过联邦学习建模，基于 10 个社区的样例数据表现，联邦学习模型性能均优于单个社区建模效果。其中两个样本数据较少的社区，联邦模型 accuracy 高于单独模型约 3%。

行业前景

基于联邦学习的智慧安防能实现保护数据安全基础上，联合多社区用户通行等数据实现对公共场所和治安防控区域全天候监控，实现提前预判、及时发现、预警和事后追溯，全方面增加社区治安防控能力，为社区平安、公安警务以及社会公共管理提供可靠有力的保障。

5.5 联邦辅助诊断

背景及需求

医疗，教育，养老，是国人最为关注的三大民生问题。近些年，由于医疗资源紧张，医生收入水平与工作强度难以匹配，医患纠纷屡屡发生。同时，分级医疗体系下，不同级别医院的医疗水平差异巨大，既有以华西医院、郑州大学第一附属医院为代表的巨无霸医院，也有广大门可罗雀的乡镇卫生所、社区医院。实力较强的医院有能力提供更好的基础设施，为医院工作人员提供有竞争力的薪资，从而进一步吸引更多的病例，进一步提升医疗科研水平；实力较弱的医院无法积累足够的病例，因而对优秀医护人员的吸引力较低，优秀医护人员的缺乏使得患者的就诊意愿进一步降低，形成恶性循环。医疗这一重要民生领域中强者愈强、弱者愈弱的发展趋势，会降低人民的幸福指数。

随着医疗电子化，以及病例及就诊信息的信息化，基于大数据的人工智能医疗成为可能。如人工智能已经在基于眼底照相的糖尿病视网膜病变等疾病检测上达到专业医生水平 [27]。这些医疗信息涉及到大量患者隐私，保护这些高度敏感信息是医院、人工智能公司、相应监管机构等各方共同的责任。

基于联邦学习的解决方案

中国的医院分为三级十等，不同医院累计的患者案例数量差异巨大。患者确诊前进行的一些标准化的化验，检验结果受到操作人员影响较小，可以依靠设备及标准化流程进行规范化。通过这些标准化的数据，引入联邦学习，利用患者健康信息，及化验、检查信息，建立横向联邦学习模型，保证患者数据仅留在就诊医院的信息系统内，帮助医生提高判断患者罹患相应疾病的准确率。

基于联邦学习的智慧医疗，不仅能够赋能实力较弱的医院，提供更高质量的检测结果，吸引更多的患者，也可以辅助医生诊断，减少医生的工作负担。同时可以将患者留在本地，减少中小城市患者远赴省城或省外就医产生的额外交通、住宿等大量成本，将资金更好地用在治疗中，减轻患者及家属负担。

联邦学习解决方案效益

以脑卒中检测为例，通过引入横向联邦学习，在病例量较小的医院，相比仅用该医院病例作为训练样本，检测准确率提升 10%~20%。随着更多的医院的病例样本加入联邦学习训练，模型的准确率将进一步提高。

同时，在基于联邦学习的疾病预测中，患者没有信息资料泄露的风险。以我国每年近 200 万的异地就医（严格限制为跨省就医）患者为例，如基于联邦学习的疾病预测若能覆盖其中百分之十人群，将在前期的确诊阶段每年节约大约 2 亿元费用。

行业前景

随着人工智能进入医疗行业，医疗机构通过电子化、信息化、结构化医疗数据，积累更多更高质量的医疗数据，为智慧医疗打下了良好的基础。同时，医疗数据涉及大量个人隐私，基于联邦学习的智慧医疗，将在保护患者隐私的基础上，赋能临床诊断等细分领域下的智慧医疗。联邦学习在智慧医疗中的应用，将以极低的成本促进优质医疗资源下沉，提升中西部地区和基层医疗机构的医疗服务能力和水平。

5.6 隐私保护广告

背景及需求

在线广告产业上下游包括广告需求方、媒介方以及广告信息交换平台。在很长一段时间内，大量互联网头部公司在实践中获得最大规模化营收的主要来源即是在线广告。国外诸如 Google，Facebook，国内如腾讯，阿里巴巴，今日头条，百度等公司，广告营收均占据较高的份额。在需求端，各个公司或品牌方每年投入大量的预算在线上广告中，来达到推广公司产品，获取未来营收等目的，如银行希望在网络上找到本行信用卡产品的潜在用户群，并通过广告争取这批潜在用户。当宏观经济增速放缓，商家的营收能力开始下滑时，将对应地削减广告预算，同时要求更高的 ROI（Return on Investment），对具有更高效率、更低成本的智慧广告提出了迫切的需求。

基于联邦学习的解决方案

线上广告核心技术包括用户定向及竞价等。基于联邦学习的广告技术，将有助于降低用户拉新的成本。利用联邦迁移学习，能够更好地整合多方数据，构建用户洞察与定向策略。在广告中重要的技术环节 RTA（Real-time API）中，有助于广告主高效获客，同时避免重复投放的资源浪费。通过在 RTA 中引入联邦技术，有助于进一步提高双方数据的安全性，同时有效降低成本。在安全性方面，通过联邦学习技术，广告主与媒体双方的用户特征与标签随机加密，双方无法知晓用户个体是否发生后端转化以及具体风险标签。同时联邦学习中差分隐私技术的接入，将混淆与打散双方数据，使对方或第三方仅能看到整体情况，而无法识别数据中的任何个人。微众银行的联邦学习联邦广告解决方案，通过预付评分，采用黑名单制，有效过滤无效流量，以降低后端转化成本，提升 ROI。

联邦学习解决方案效益

在贷款、信用卡、保险、在线教育、家装行业等具有高价值行业的广告产品中，相比于传统广告仅优化前端数据的解决方案，联邦广告技术通过安全使用加密的合作方后端数据，在传统的展现、点击、到达环节外，涵盖了转化端的数据，同时严格保证了用户及平台的数据隐私。其中，以联邦广告技术针对金融行业的前后端打通方案为例，在贷前阶段，反作弊转化效果提升 2%，同时通过 RTA 直接实现，不需要客户进行任何操作，在贷后阶段，通过流量分级，过滤高风险无法通过授信的用户，使转化效果提升 5% 左右。

行业前景

以联邦学习，差分隐私和多方安全计算为保障，实现转化数据安全与用户隐私的联邦广告优化技术，有助于防控风险流量，从而针对线上流量进行优化，提前过滤无效流量，同时在日益趋严的隐私政策监管下，保护数据隐私安全并优化投放效率。通过联邦广告技术，减少投放产本，提升广告主 ROI，将资金更好地用于产品的创新与研发中；同时对于广告供给方，可以提高访问用户点击率及转化率，优化链路效率，达到多方共赢。

5.7 联邦自动驾驶

背景及需求

我国的城市内道路场景复杂多样，给无人驾驶带来了严峻的挑战。但以高速公路，码头，无人园区为代表的典型交通场景，由于道路环境简单，行人及车辆的突发状况较少，仍然是适宜无人驾驶落地的重要场景。其中物流等行业每年需要大量的司机长距离行驶在高速公路上，愿意从事这一行业的年轻人越来越少。同时，高发的道路安全事故也给这个行业笼罩上一层阴影。据报道，卡车司机中每 7 个人就有一人发生过交通事故，而卡车司机每年的死亡率高达 5‰[28]。无人驾驶将使人从高压且高危的高速公路司机行业中解放。同时，无人卡车运输可以支持 7 天 24 小时运行，提升了资产利用率，降低了成本。

随着 Google 无人车项目 Waymo[29] 上线，特斯拉推出无人驾驶系统 Autopilot[30]，百度推出 Apollo[31]，Uber、滴滴等共享出行头部公司布局无人驾驶，无人驾驶达到了空前的热度。伴随着各大车企开始通过以辅助驾驶为亮点的差异化竞争，车联网，路联网技术的推演，未来无人驾驶依然是一个具有极高社会价值的和经济价值的技术方向。

基于联邦学习的解决方案

一个普通的车辆受制于驾驶的时间和空间限制，通常获取到的传感器信息是有一定局限的。通过引入横向联邦学习，融合不同车辆的摄像头、超声波传感器、雷达（如毫米波雷达、激光雷达）等传感器信息，可以更加快速地建立场景信息，同时有助于提高模型的鲁棒性。

近些年，人们逐渐意识到，无人驾驶不应该只是简单学习或复制人类个体的驾驶能力，还可以与车联网、车路协同，甚至整个交通系统共同交互，来创造更好的驾驶环境。车辆与系统环境的交互学习，可以辅助以城市的其它信息，诸如城市摄像头、交通灯、未来的智能道路等，通过纵向联邦的方式来更好地在隐私保护下融合不同来源信息，提升无人驾驶体验。

联邦学习解决方案效益

感知、规划（决策）、控制是无人驾驶的三个核心模块。其中，以激光雷达、相机（单目、双目、环视摄像头）等传感器信息输入的感知，类似于人的眼睛，为后续规划阶段提供基础的输入，是无人驾驶的基础。单车产生的数据通常具有时间和空间的局限性。同时，由于车辆行驶过程中将产生海量传感器数据，原始数据可能涉及到隐私问题，使得多车辆的数据在传统中心式机器学习下，可能带来新的伦理与技术挑战。

利用不同车辆的数据，同时保护数据隐私，减少通信带宽，通过联邦学习建模，基于 nvidia-Jetson RC 实验车，联邦学习在避障、道路规划等项目上性能显著优于单车的学习。其中在避障子项上，基于联邦学习的实验车，避障性能优于普通实验车 48% 以上。

行业前景

汽车产业是国民经济的支柱产业，在新的十年里，迫切需要新的技术带来新的产业增长点。2018 年，我国劳动力人口出现首次下降（引自国家统计局数据），伴随着老龄化的加速，以及社会经济的持续发展，未来运输行业的劳动人口成本将不断提高。从事运输行业的高危性质及有限的发展空间，将更加难以吸引年轻人从事该行业。无人驾驶技术的发展是未来二十年促进社会良性发展的关键技术之一。基于横向联邦学习的无人驾驶技术，有利于在保护司乘用户的隐私前提下，加速对环境感知的学习能力。在未来，基于纵向联邦学习的无人驾驶，辅助物联网，车路协同，5G 等新技术，将建设一个高效、安全、低成本的智慧交通环境。

第六章
联邦学习的发展路径

结合人工智能与大数据的发展环境，行业痛点与需求的实际情况，建议通过以下四个阶段来发展联邦学习：

6.1 培育联邦学习开源发展生态

建立良好的开源环境是联邦学习可持续性发展的基础。良好的开源软件支撑，有利于扩大联邦学习技术的影响力，吸引更多参与方投身联邦学习的技术研究与业务落地，进一步使整个产业上下游联动起来，积极参与联邦学习技术开源工作。

开源生态环境的培育可通过以下方面考虑：

开源社区的管理与运作，需要较高的资本与人力投入，所以大部分是依托公司、基金会、联盟成立的。将联邦学习开源框架依托于如 Github，Linux 基金会，Apache 基金会等，有利于在短期内吸引全球相关人员进行快速地使用。在实际应用中，全球不同公司，机构或个人反馈第一手建议，开发者根据建议为联邦学习制定高效实用的短期和长期研究目标。同时，将联邦学习平台托管于开源社区，能吸引大量人员加入开源框架发展，与维护者共同开发，促进联邦学习生态整体形成良好的正反馈。

建立更加开放，鼓励公平竞争的国内开源环境是联邦学习自主应用的基础。我国作为最有潜力的发展中国家，具有庞大的潜在市场，及为数众多的高素质科研、技术人员，培育国内良好的联邦学习开源环境，将有利于相关人员更加高效地进行联邦学习技术的交流与应用，同时减少对全球政治经济环境波动时期下对技术管控的依赖。目前集成联邦学习计算工具集的 OpenI 纵横，作为一个完成的项目捐献给 OpenI 启智平台。OpenI 纵横提供了丰富的一站式联邦建模算法组件，方便快速实验和迭代算法，满足大多数联邦建模任务。

6.2 建立联邦学习国内外标准

国内、国际上都在加速人工智能标准体系和相关标准的建设。国际标准化组织于 2017 年 10 月成立人工智能分技术委员会（ISO/IEC JTC 1/SC 42），美国、德国等国家提出了人工智能术语、参考模型标准项目提案。国内于 2018 年 1 月，在国标委和工信部的高度重视和指导支持下，成立了国家人工智能标准化总体组，集合国内在人工智能领域的重要企业和科研院所，推动我国人工智能标准体系的建设。

通过研制和建立联邦学习的国内标准（如团体标准和国家标准）与国际标准（如 IEEE 企业标准），制定联邦学习的算法框架规范，使用模式和使用规范，可帮助不同类别的企业在合作过程中合法合规的共同使用数据，

在用户的隐私和数据安全的情况下，不同的数据实体合作共赢，建立更准确的数据模型。也给人工智能在不同产业的实际落地中或将遇到的问题，提供可行性依据。中国通信标准化协会大数据技术标准推进委员会（TC601）也正在开展联邦学习相关的行业标准和团体标准的制定，并即将启动联邦学习类产品的标准符合性测试。行业方面，由中国人民银行科技司牵头制定的金标委行业标准《多方安全计算金融应用技术规范》，当前已进入征求意见阶段；今年，互联网金融协会也将在此框架之下，启动联邦学习数据合作规范相关的团体标准制定工作，强调可落地性，针对不同场景的数据合作防护等级要求，提出参考性实现。

6.3 建立行业垂直领域应用示例

联邦学习在产业场景中的实际落地，将给算法研究提供切实有效的支撑。应用场景可分为同构场景和异构场景。同构场景指的是两个企业属于相同或相近的领域，所拥有的数据性质相似，特征相近，但是样本不同。如在银行和金融机构间的合作，双方拥有的不同的用户样本，但是样本属性同质，这种场景下使用横向联邦学习，可达到将双方样本放到一起的建模效果。异构场景指的是两个企业分属不同的领域，所拥有的数据性质不同，特征不同，但是有重叠的样本 ID。比如银行与互联网公司之间的合作，双方有重叠的用户 ID，但是企业间各自拥有用户不同的特征，如银行有用户的收入和交易行为，互联网公司有用户的社交或出行行为，这种场景下使用纵向联邦学习建模，可达到特征增加的建模效果。两种场景下的应用均可使得比数据在本地单方建模更好。

推动联邦学习在行业垂直领域的应用尤其是异构场景下的应用，将建立一个基于联邦学习的新的数据商业模式和共同成长的大数据生态。

6.4 全面展开建立联邦数据联盟

联邦学习的发展大体需要经历三个阶段：一阶段是点到点的联邦学习发展阶段，包括发展开源联邦，建立联邦标准；二阶段是应用落地、积累案例阶段；第三阶段是建立联邦数据联盟阶段。联邦学习的期望是把数据背后的真正的知识和价值拿出来，参与各方共建一个联邦数据联盟。在联邦数据联盟中，已经不再局限于传统基于知识图谱的知识价值网络，而是一张张拥有切实可行的数据价值及行业知识的大图，通过激励机制鼓励联邦联盟中的已有成员，让他们各自的数据在合法合规下带来真正的价值流动，为自身带来收益，同时吸引该领域内更多公司或机构加入成为新的节点，让节点之间的连接更加广泛，产生有价值的动态流动。联邦数据联盟的成立，使数

据成为杠杆，撬动行业利润以低成本高速增长，激励机制使垂直领域的连接更加紧密，联邦学习技术保证行业良性发展。千行万业建立各自的联邦数据网络，金融业拥有金融联邦数据网络，医疗业拥有医疗联邦数据网络，零售业拥有零售数据网络，以及运输业，快递业，旅游业等等。在未来，不同的网络间还将有所交互，诞生无穷的想象空间。

第七章
总结展望

近年来，数据的孤岛分布以及对数据隐私监管力度的加强正在逐渐成为人工智能的下一个挑战，联邦学习的产生为人工智能打破数据屏障和进一步发展提供了新的思路。它实现了在保护本地数据的前提下让多个数据拥有方联合建立共有的模型，从而实现了以保护隐私和数据安全为前提的互利共赢。本文概括性地介绍了联邦学习的基本概念、构架与技术原理，并且尝试在一些应用场景中探讨联邦学习对人工智能发展的巨大助力。期待在不远的将来，联邦学习能够帮助打破各领域、各行业的数据壁垒，在保护数据隐私和安全的前提下形成一个数据与知识共享的共同体，并同时解决了奖励对联盟做出贡献机构的共识机制，必将能为人工智能带来的红利落实到社会的各个角落。

7.1 联邦学习的未来研究方向

7.1.1 安全性

联邦学习中，以下部分可能遭受到攻击：

客户端：对客户端设备具有管理员访问权限的人，可以通过控制客户端进行恶意攻击。被恶意操控的客户端可以在它们参与的迭代中，检查从服务器接收到的所有消息（包括模型），并可以篡改训练过程。中立的客户端可以检查从服务器收到的所有消息，但不会篡改训练过程。

服务端：被恶意操控的服务器可以在所有迭代中检查发送到服务器的所有消息（包括梯度更新），并可以篡改训练过程。中立的服务器可以检查发送到服务器的所有消息，但不会篡改培训过程。

同时，在模型输出以及部署的过程中，也可能遭受到恶意攻击。在这种情况下如何严格保证隐私，是一个极大的挑战。

从攻击手段上看，主要有模型更新攻击（model update poisoning），数据攻击（data poisoning attack）和逃逸攻击（evasion attack）。依据在生命周期中的位置，这些攻击也可大致分为训练时间攻击（模型更新攻击，数据攻击）和推理时间攻击（逃逸攻击）。

模型更新攻击：恶意攻击者可以直接控制一些客户端，并改变这些客户端的输出，从而使所学习的模型偏向于他们的目标。当恶意攻击者可以控制客户端产生任意输出时，这种攻击称为拜占庭攻击 [32]。该攻击下，受控的客户端可以发送任意值，而不是将本地更新的模型发送到服务器。这可能导致收敛到次优模型，甚至导致模型发散。相比拜占庭攻击式的无目标攻击，有目标模型攻击通常需要较低的成本。文献 [33] 显示，在联邦学习中，当 10% 的设备被恶意者控制，即有可能通过攻击服务器的模型引入后门。在中心化机器学习中，通过控制训练

过程的方法减少模型更新攻击的防护手段，在联邦学习中无法直接应用。

数据攻击：区别于模型更新攻击，在数据攻击中，恶意攻击方不能直接更改中心节点的模型，但可以通过篡改客户端的数据，特征或者标签达到无目标攻击或者针对特定目标攻击的目的。与模型更新攻击相同，仅靠全局准确率或单客户端训练准确率等指标较难检测恶意的数据攻击的存在。

逃逸攻击：在模型推理阶段，攻击者可以在不改变机器学习系统的情况下，通过构造特定输入样本以完成欺骗目标系统的攻击。通过增加噪声等方式，产生在人类看来与原始的测试输入几乎没有区别的输入，却可以欺骗经过训练的模型。在图像和音频领域，对抗样本通常是通过在测试样本中加入范数有界的扰动来构建的。对抗性训练（用对抗性样本训练一个健壮的模型）通常对白盒规避攻击具有一定的健壮性，然而对抗性训练通常只会提高对训练中包含的对抗性样本这种特定类型样本的健壮性，训练后的模型依然容易受到其他形式的对抗性噪声的影响。同时，通过对抗训练来减少逃逸攻击的方法，在联邦学习中可能存在以下问题：对抗性训练主要是针对独立同分布数据开发的，而在非独立同分布环境中它的表现并不清楚；在无法在训练前检查训练数据的联邦学习中，较难设置适当的扰动范数界限。因此，在联合学习设置中可能需要新的鲁棒优化技术来解决逃逸攻击。

差分隐私等技术是减小攻击的一个主要技术。联邦学习系统中的许多挑战可以被看作是确保一定程度的健壮性：不管是否是恶意的，干净的数据被破坏或以其他方式篡改。差分隐私 (DP)[34] 从健壮性的角度定义了隐私。简而言之，通过在训练或测试时加入随机噪声，以减少特定数据点的影响。

除了恶意的攻击外，与传统的中心化机器学习相比，联邦学习也可能受到来自服务提供者控制之外的不可靠客户端的非恶意故障的影响。虽然非恶意的失败通常比恶意攻击的破坏性更小，但它们可能更常见，并且与恶意攻击具有共同的根源和复杂性。因此，未来在安全方面的研究，不仅包含防范恶意攻击，也包括减少非恶意的故障带来的隐私安全影响。

7.1.2 激励机制

联邦学习的价值在于打破数据孤岛，通过鼓励具有相同数据结构（横向联邦学习）或不同数据结构（纵向联邦学习）共同参与训练，提高模型的整体效果。在整个过程中，一个有效的激励机制的设计，可以激励更多终端用户，或者不同企业、组织参与联邦学习。博弈论、契约理论等的引入，可以更好地帮助设计激励机制。同时，在不同终端或组织参与的过程中，需要有效衡量不同参与方的贡献程度，从而根据贡献程度公平地分配奖励给参与方，进一步提高用户或组织的贡献热情，形成一个良好的正循环。一个联邦学习下的有效的奖惩及分配机制的设计，也有其重要研究价值。

7.1.3 有效性和效率

非独立同分布的数据

现实世界中，大量的数据分布是非独立同分布 (non-IID) 的，例如：不同地理区域的人有不同的喜好与倾向；在特殊的时间段下，一个人会做出与平时不同的差异选择；不同的客户端拥有的数据量大小可能有巨大的差异等。

中心化机器学习可以获取全部样本或者过去已产生的全部样本，从而完成全局最优的模型训练。在联邦学习中，由于数据无法出本地的限制，传统的中心化机器学习中大量调参方法，如随机化数据顺序（data shuffle）无法被直接应用。相比中心化的机器学习训练过程，这种数据分布造成的影响将带来训练模型效果的降低。

通过对目标函数进行改进，或者进行有限假设是减少非独立同分布数据影响的一个研究方向。为了解决数据分布造成的影响，FedProx[35] 算法提出在每个局部目标函数中加入一个近端项，使算法对局部目标的不均匀性更加鲁棒。Ahmed Khaled[36] 假设所有客户端都参与，并对客户端使用批梯度下降，该方法下客户端上的随机梯度收敛得更快。

通过对优化函数进行改进也是解决数据非独立同分布的一个研究方向。在深度学习中，优化函数经历了 SGD，Adagrad，Adam 等发展，其中动量概念的引入带来了更快的收敛速度与精度。在联邦学习优化函数中，对于一阶优化方法，动量和方差的引入是提高优化和泛化性能的有效途径。然而，对于如何将动量或方差技术纳入联邦学习相关的局部 SGD 和联合平均，目前还没有达成共识。SCAFFOLD[37] 使用控制变量显式地对客户端更新中的差异进行建模，以执行方差减少，这可以在不限制客户端数据分布差异的情况下快速收敛。对于动量方案，Yu 等人 [38] 建议在每个客户端维护一个本地动量缓冲区，并在每个通信回合平均这些本地缓冲区以及本地模型参数。虽然这种方法在实验中提高了本地 SGD 的精度，但它需要双倍的通信成本。Wang 等人 [39] 提出了 SlowMo 的动量方案，它可以在不牺牲吞吐量的情况下显著提高局部 SGD 的优化和泛化性能。Hsu 等人 [40] 提出类似于 SlowMo 的动量方案。局部 SGD 的动量变量可以凭借与同步小批量 SGD 相同的速度收敛到非凸目标函数的平稳点，但要证明动量加速了联邦学习环境下的收敛速度依然在理论上存在困难 [38-39]。

此外，微调（fine-tuning），迁移学习（transfer learning），元数据学习（meta learning）等技术也在不断被引入联邦学习，探索如何解决非独立同分布数据带来的影响。

有限资源下的参数调节

在联邦学习中，除了具有与深度学习或传统机器学习相似的优化函数选择，如学习率，批量大小，正则化等，还需要考虑聚合规则，每一个迭代中选择的客户端数量，本地每轮的迭代数量等参数选择。联邦学习的参与方中，可能是拥有较多计算和存储资源的数据中心服务器，也可能是不定时在线的边缘设备。一些参与方可能仅拥有有

限的计算、存储、网络资源。一些在深度学习中帮助调节模型性能的方法，诸如AutoML，NAS（neural architecture search）等，由于将占用较多的资源，将直接降低通信和计算效率，无法直接在联邦学习中应用。在有限资源下的超参数调节是一个极具挑战和有意义的研究方向。

有限的通信带宽及设备的不可靠性

通过无线网络接入，或者互联网中靠近终端的端方用户，相比数据中心或者数据中心链路上的核心节点，通常拥有较低的网络带宽及通信效率，同时这种网络连接可能有较高的花费，或者无法保证完全稳定在线。例如，一个终端手机可能仅在电量充足并连接至无线网络的情况下，接入联邦学习训练。这引发了学者对减少联邦学习的通信带宽的研究。在梯度，模型传播，局部计算等部分均有数据压缩的空间。将联邦平均与稀疏化，或者和量化模型更新结合的方法，已经证明在对训练精度影响较小的情况下显著降低了通信成本。然而，目前还不清楚是否可以进一步降低通信成本，以及这些方法或它们的组合是否能够在联邦学习中提供通信效率和模型准确性之间的最佳平衡。

参考文献

[1]新华网. (受权发布) 中华人民共和国网络安全法[OL].[2020-02-17].
http://www.xinhuanet.com/politics/2016-11/07/c_1119867015.htm
[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]. [2020-02-17].
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]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
[5] Dwork C. Differential privacy: A survey of results[C]//International Conference on Theory and Applications of Models of Computation. Springer, Berlin, Heidelberg, 2008: 1-19.
[6] Sweeney L. k-anonymity: A model for protecting privacy[J]. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2002, 10(05): 557-570.
[7] Li N, Li T, Venkatasubramanian S. t-closeness: Privacy beyond k-anonymity and l-diversity[C]//Data Engineering, 2007. ICDE 2007. IEEE 23rd International Conference on. IEEE, 2007: 106-115.
[8] Craig Gentry et al. Fully homomorphic encryption using ideal lattices. In Stoc, volume 9, pages 169–178, 2009.
[9] Zvika Brakerski. Fully homomorphic encryption without modulus switching from classical gapsvp. In CRYPTO, volume 7417 of Lecture Notes in Computer Science, pages 868–886. Springer, 2012.
[10] Junfeng Fan and Frederik Vercauteren. Somewhat practical fully homomorphic encryption. IACR Cryptology ePrint Archive, 2012:144, 2012.
[11] Zvika Brakerski, Craig Gentry, and Vinod Vaikuntanathan. (leveled) fully homomorphic encryption without bootstrapping. In ITCS, pages 309–325. ACM, 2012.
[12] Jean-S´ebastien Coron, Tancr`ede Lepoint, and Mehdi Tibouchi. Scale-invariant fully homomorphic encryption over the integers. In Public Key Cryptography, volume 8383 of Lecture Notes in Computer Science, pages 311–328. Springer, 2014.
[13]Ho Q, Cipar J, Cui H, et al. More effective distributed ml via a stale synchronous parallel parameter server[C]//Advances in neural information processing systems. 2013: 1223-1231.
[14] Sheth A P, Larson J A. Federated database systems for managing distributed, heterogeneous, and autonomous databases[J]. ACM Computing Surveys (CSUR), 1990, 22(3): 183-236.

参考文献

[15]Konečný J, McMahan H B, Yu F X, et al. Federated learning: Strategies for improving communication efficiency[J]. arXiv preprint arXiv:1610.05492, 2016.

[16]Bonawitz K, Ivanov V, Kreuter B, et al. Practical secure aggregation for federated learning on user-held data[J]. arXiv preprint arXiv:1611.04482, 2016.

[17]Truex S, Baracaldo N, Anwar A, et al. A hybrid approach to privacy-preserving federated learning[J]. arXiv preprint arXiv:1812.03224, 2018.

[18]Liu Y, Chen T, Yang Q. Secure Federated Transfer Learning[J]. arXiv preprint arXiv:1812.03337, 2018.

[19] 微众银行. 联邦学习开源平台FATE [CP/OL]. [2020-02-17]. https://github.com/FederatedAI/FATE

[20] McMahan H B, Moore E, Ramage D, et al. Communication-efficient learning of deep networks from decentralized data[J]. arXiv preprint arXiv:1602.05629, 2016.

[21] Pan S J, Yang Q. A survey on transfer learning[J]. IEEE Transactions on knowledge and data engineering, 2010, 22(10): 1345-1359.

[22] Liu Y, Chen T, Yang Q. Secure Federated Transfer Learning[J]. arXiv preprint, arXiv: 1812.03337, 2018.

[23] 微众银行. 工业级联邦学习框架 [OL]. [2020-02-17]. https://fate.fedai.org/

[24]谷歌. TensorFlow Federated：基于分散式数据的机器学习[OL]. [2020-02-17]. https://tensorflow.google.cn/federated

[25]百度飞桨. 百度飞桨release note[OL]. [2020-02-17]. https://www.paddlepaddle.org.cn/

[26]Openmined.syft[OL]. [2020-02-17]. https://pypi.org/project/syft/

[27] Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs[J]. Jama, 2016, 316(22): 2402-2410.

[28]新浪财经. 卡车司机生存纪实：留下来的边缘人每天都是穷途末路[OL]. [2020-02-17]. http://finance.sina.com.cn/china/gncj/2018-10-08/doc-ifxeuwws2145661.shtml

[29] Waymo. Waymo官网[OL]. [2020-02-17]. https://waymo.com

[30]特斯拉.Autopilot系统介绍[OL].[2020-02-17]. https://www.tesla.cn/autopilot

[31]百度.apollo自动驾驶解决方案[OL].[2020-02-17]. http://apollo.auto/

[32] Lamport L, Shostak R, Pease M. The Byzantine generals problem[M]//Concurrency: the Works of Leslie Lamport. 2019: 203-226.

参考文献

[33] Bhagoji A N, Chakraborty S, Mittal P, et al. Analyzing federated learning through an adversarial lens[J]. arXiv preprint arXiv:1811.12470, 2018.

[34] Dwork C, McSherry F, Nissim K, et al. Calibrating noise to sensitivity in private data analysis[C]//Theory of cryptography conference. Springer, Berlin, Heidelberg, 2006: 265-284.

[35] Li T, Sahu A K, Zaheer M, et al. Federated optimization in heterogeneous networks[J]. arXiv preprint arXiv:1812.06127, 2018.

[36] Khaled A, Mishchenko K, Richtárik P. First analysis of local gd on heterogeneous data[J]. arXiv preprint arXiv:1909.04715, 2019.

[37] Karimireddy S P, Kale S, Mohri M, et al. SCAFFOLD: Stochastic controlled averaging for on-device federated learning[J]. arXiv preprint arXiv:1910.06378, 2019.

[38] Yu H, Jin R, Yang S. On the linear speedup analysis of communication efficient momentum sgd for distributed non-convex optimization[J]. arXiv preprint arXiv:1905.03817, 2019.

[39] Wang J, Tantia V, Ballas N, et al. SlowMo: Improving Communication-Efficient Distributed SGD with Slow Momentum[J]. arXiv preprint arXiv:1910.00643, 2019.

[40] Hsu T M H, Qi H, Brown M. Measuring the effects of non-identical data distribution for federated visual classification[J]. arXiv preprint arXiv:1909.06335, 2019.

WeBank 微众银行 | FedAI

官网：https://www.fedai.org

邮箱：contact@fedai.org

FATE GitHub：https://github.com/FederatedAI/FATE

FATE开源社区公众号

电子商务电子支付国家工程实验室公众号

鹏城实验室公众号

平安科技公众号

腾讯研究院公众号

大数据技术标准推进委员会公众号

招商金融科技公众号