小 型 微 型 计 算 机 系 统
Journal of Chinese Computer Systems
DOI:10.20009/j.cnki.21-1106/TP.2025-0399
2026 年 5 月 第 5 期
Vol.47 No.5 2026

个性化联邦学习综述

王 鑫[1,2],黄启超[1],孙凌云[2]

[1](浙江工业大学 计算机科学与技术学院,杭州 310023)
[2](浙江大学 计算机科学与技术学院,杭州 310058)
E-mail:xinw@zjut.edu.cn

摘 要:联邦学习(Federated Learning,FL)是一种典型的分布式机器学习方法.由于在隐私保护方面有着独特的优势,联邦学习在近年来受到了广泛的关注和研究.然而,传统的联邦学习存在着两大亟待解决的核心难题:数据异构性困境、模型泛化与个性化之间的冲突.为了解决这些问题,个性化联邦学习(Personalized Federated Learning,PFL)的概念被引入,它能够针对每个联邦学习客户端的本地数据特点进行个性化的模型调整,允许联邦学习客户端在保护自身敏感数据隐私的同时根据自己的需求构建个性化模型.本文概述了个性化联邦学习的概念以及目前所面临的关键问题,分类综述了个性化联邦学习各类方法的发展现状,同时介绍了个性化联邦学习的一些新兴研究方向和领域应用情况.

关 键 词:联邦学习;个性化联邦学习;知识蒸馏;预训练-微调;个性化特征聚合
中图分类号:TP181 **文献标识码**:A **文 章 编 号**:1000-1220(2026)05-1117-10

Survey of Personalized Federated Learning

WANG Xin[1,2],HUANG Qichao[1],SUN Lingyun[2]

[1](College of Computer Science and Technology,Zhejiang University of Technology,Hangzhou 310023,China)
[2](College of Computer Science and Technology,Zhejiang University,Hangzhou 310058,China)

Abstract:Federated Learning (FL) is a typical distributed machine learning approach. Due to its unique advantages in privacy preservation,FL has garnered widespread attention and research in recent years. However,traditional federated learning faces two core challenges that urgently need to be addressed:the issue of data heterogeneity and the conflict between model generalization and personalization. To tackle these problems,the concept of Personalized Federated Learning (PFL) has been introduced. PFL enables personalized model adjustments based on the local data characteristics of each federated learning client,allowing clients to build customized models according to their specific needs while protecting the privacy of their sensitive data. This paper provides an overview of the concept of personalized federated learning and the key challenges it currently faces,categorizes and reviews the developmental progress of various PFL methods,and introduces some emerging research directions and practical applications of personalized federated learning.

Keywords:federated learning; personalized federated learning; knowledge distillation; pre-training and fine-tuning; personalized feature aggregation

0 引 言

联邦学习(Federated Learning,FL)是一种创新的分布式机器学习方法,是在异构边缘计算环境下的一种分布式学习范式[1,2],其核心设计理念为打破数据孤岛,实现多方数据协同利用,同时最大程度保障数据隐私安全.FL 颠覆了传统机器学习对集中式数据收集的依赖模式,允许多个参与方在不直接共享原始数据的前提下共同参与模型训练[3].个性化联邦学习(Personalized Federated Learning,PFL)是在 FL 基础上发展而来的前沿技术,它保留了 FL 的各项特性,但更加侧重于训练特定于 FL 客户端的模型,寻求通过协作学习为每个客户量身定制个性化的模型,在每个 FL 客户端上实现令人满意的个性化性能,而不是使用统一的全局模型来适应所有 FL 客户端的数据[4-6].

在实际场景中,各参与主体即便处于同一领域,其内部的数据分布和业务目标等也都存在差异.PFL 作为一种隐私保护技术,借助加密技术和安全协议来保障数据交互的安全,并以参数或梯度等中间结果的传递来推动模型训练.在这个过程中,各参与方既能依据 FL 框架参与全局模型训练,又能依据自身独特的本地数据去统计特征表示和个性化诉求,灵活地对全局模型进行个性化微调.PFL 通常一方面利用模型压缩和迁移学习策略帮助 FL 客户端快速适配全局模型,另一方面通过个性化层架构巧妙地融合全局共享知识与本地个性化特征,从而打造出贴合 FL 客户端需求的专属模型.本文研究 PFL,主要是因为它对于模型个性化有着非常大的帮助.各类智能设备和互联网应用的普及产生了海量的数据,这些数据的背后是一个个具有独特行为模式、兴趣偏好和需求的用户,众多软件和数据分析系统都急需能够精准匹配每个用户

收稿日期:2025-10-16 收修改稿日期:2025-12-15 基金项目:浙江工业大学科技项目(KYY-HX-20220288,KYY-HX-20180649)资助. 作者简介:**王 鑫**(通信作者),男,1984 年生,博士,副教授,CCF 会员,研究方向为机器学习、大数据分析、联邦学习等人工智能技术;**黄启超**,男,2002 年生,硕士研究生,研究方向为个性化联邦学习;**孙凌云**,男,1981 年生,博士,教授,博士生导师,CCF 会员,研究方向为人工智能、设计智能、信息与交互设计.

或业务场景的个性化模型来契合用户的多元需求、提升用户

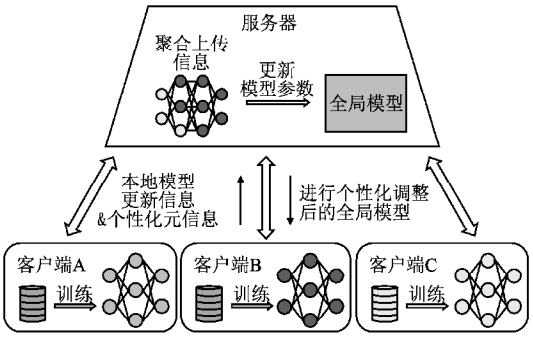

图 1 个性化联邦学习典型架构

Fig. 1 Typical architecture of personalized federated learning

体验[7]. 图 1 为 PFL 的典型架构.

1 关键问题

PFL 领域尚有许多亟待解决的问题,其中有一些是这个研究领域当中众多学者公认的关键问题和难题,比如数据异构性问题、模型融合问题和本地与全局模型的平衡问题. 本节会对这 3 个问题进行详细的解读.

1.1 数据异构性问题

数据异构性,从广义上讲,是指参与 FL 或 PFL 的各个客户端所持有的数据在多维度上呈现出的显著差异特性,其对于 PFL 有着非常严重的影响[8]. Ye 等人[9]将数据异构性挑战分为 4 个部分:标签偏移、特征偏移、质量偏移 和数量偏移. 在 FL 迭代训练过程中,数据异构性会使不同客户端计算出的模型参数更新方向与步长差异巨大. 当服务器聚合这些参数时,容易引发更新冲突,导致模型难以朝着全局最优方向收敛,训练过程的反复震荡会耗费大量计算资源与时间成本. 在这个过程中,模型学习到的知识局限于各客户端的本地数据模式,无法有效提炼出适用于更广泛场景的通用规律. 另外,数据异构性会使客户端基于本地数据对全局模型的个性化调整难度倍增,不仅全局模型难以兼容各类异构数据特征,个性化层与全局模型的融合效果也会受到干扰,最后导致个性化模型偏离实际需求[10].

1.2 模型融合问题

在 PFL 架构下,各参与方需要整合基于自身独特数据训练出的个性化模型,但各方的模型结构存在差异,参数特性各不相同,且模型在融合过程中还需权衡个性化与通用性,这些因素都会使模型在融合时难以找到统一的规则,导致融合后的模型无法有效兼顾各方需求,很难达到理想性能. 在实际应用当中,为了满足用户多样化的使用习惯和场景需求,不同设备训练出的个性化模型需要频繁融合. 这样频繁的融合操作,一方面会让融合算法的计算复杂度急剧上升,消耗大量的计算资源,另一方面也使得融合过程中出现错误和冲突的概率增加. 同时在对模型性能要求极高的场景中,因为无法实现有效的模型融合,不能充分利用各参与方的个性化数据,这使得 PFL 难以在这些场景展现其价值.

1.3 本地与全局模型的平衡问题

本地与全局模型的平衡问题,核心在于如何在 FL 框架下,协调好各参与方对本地个性化模型的需求与全局共享模型目标达成之间的矛盾关系. 个性化模型需要依据各客户端独特的本地数据特征进行定制,常常额外引入适配本地特殊情况的模块. 当与全局模型融合时,新增的个性化架构层可能会破坏全局模型已学习到的共性知识. 在制定训练策略时,若侧重于全局模型训练,就容易忽略各客户端数据的细微差异,使得客户端在应用全局模型时本地个性化需求得不到满足,模型对本地数据的拟合效果差;若过度倾向个性化训练,模型容易陷入过拟合状态,泛化能力下降,各客户端模型逐渐偏离,无法再有效共享知识. 在数据利用环节,全局模型训练需要汇聚各方数据的共性特征信息,在一定程度上会对数据进行抽象,弱化个体数据细节,但个性化模型恰恰依赖于这些本地数据的独特细节. 如何在为全局模型提供充足共性数据的同时保留客户端个性化数据的价值是必须要考虑的一环. 表 1 展示了不同训练策略的优点和缺点.

表 1 不同训练策略的优缺点

Table 1 Advantages and disadvantages of different training strategies

训练策略	优点	缺点
侧重本地模型(如[11,12])	满足本地个性化需求,数据拟合效果好	可能破坏全局共性知识;易过拟合,削弱协同性;
侧重全局模型(如[13])	捕捉数据共性,构建泛化强的模型	忽略本地差异,个性化不足;弱化了数据细节
兼顾本地与全局(如[1,3,14])	平衡本地个性化与全局协同,提升整体性能	协调难度大;策略易失衡;数据平衡困难

2 个性化联邦学习方法

为了系统性地梳理 PFL 的技术脉络,本文根据现有方法在实现个性化过程中所遵循的核心技术路线,将其划分为 7 个类别. 不同的 PFL 方法被分到哪一类别,则主要依据其在实现客户端模型个性化过程中所采用的根本性机制. 具体来说:1)基于知识蒸馏的方法,其核心是通过师生模型间的知识迁移来实现个性化;2)基于预训练-微调的方法,其核心是采用预训练模型 + 本地微调适配的两阶段范式;3)基于个性化特征聚合的方法,其核心在于改进服务器端的聚合策略;4)基于元学习的方法,其核心是优化模型使其具备快速适应新客户端的能力;5)基于模型正则化的方法,其核心是在客户端本地训练目标中引入正则化项来约束本地更新与全局信息的一致性;6)基于专家混合模型的方法,其核心是为不同数据分布或任务类型的客户端动态选择或组合多个专家模型. 上述 6 个类别基本涵盖了当前 PFL 领域的主流技术范式. 此外,部分研究的出发点或技术思路较为独特,难以纳入以上类别当中,因此本文将它们归为其他方法并单独讨论,以确保分类框架的完备性.

2.1 基于知识蒸馏

知识蒸馏技术是一种模型压缩方法,通过将知识从预先训练过的教师模型转移到学生模型,即利用一个复杂且性能良好的教师模型学到的知识来训练相对简单的学生模型,达到提高学生模型性能的目的[3,15]. 知识蒸馏技术在深度学习领域,特别是模型需要部署在资源受限设备上时可以发挥良

好的作用[16].

知识蒸馏可以作为解决 FL 中数据异质性问题的实用解决方案[3],其中的自知识蒸馏方法被研究者所广泛关注. 例如 pFedSD(PFL via self-knowledge distillation)[1],通过本地自知识蒸馏的方式显著提高了算法的个性化性能和收敛性,旨在解决边缘场景中的统计异构性挑战. 在框架中,FL 客户端进行本地训练,并保存更新后的模型作为下轮教师模型,最终由 FL 服务器聚合生成新的全局模型,通过自知识蒸馏机制防止 FL 客户端遗忘历史个性化知识,实现个性化与泛化的动态平衡. 另一边,FedBSD(PFL via Backbone Self-Distillation)[3]作为基于主干网络的自知识蒸馏方法,同样引入了一个自蒸馏框架. FL 客户端在训练完本地模型后仅上传主干权重至服务器,用于聚合生成全局主干,FL 客户端将收到的全局主干与本地主干分别作为教师和学生模型进行知识蒸馏,在更新本地主干的同时实现全局知识融合和整体性能的提升,减轻了传统知识蒸馏方法对于模型个性化程度的影响. 两种方法各有特点,pFedSD 与其他算法的兼容性好,而 FedBSD 被证明其在实际应用场景中具有高度适用性,不过作为自知识蒸馏方法,他们都会增加 FL 客户端的计算负担. 除此之外,Peng 等人[17]提出的方法 PSFL(Personalized Semantic Federated Learning)虽然不是自知识蒸馏方法,但也结合了一种个性化本地蒸馏策略和自适应全局剪枝方法. 个性化本地蒸馏策略允许 FL 客户端根据资源选择不同复杂度的教师模型,并通过学生模型实现联邦聚合的统一性,自适应全局剪枝方法则会实时动态调整模型剪枝比例,二者结合实现了模型个性化与通信效率的双重优化. 但同样,该方法的计算开销较大,同时训练教师和学生模型会增加更多的客户端计算成本,导致资源受限设备可能无法承受.

一些研究者在现有的知识蒸馏方法上进行了一定程度的创新. 为了更有效地区分通用模型表征与个性化模型表征之间的相似性及差异性,Chen 等人[18]首创了一种基于傅里叶频谱的模型相似性度量方法. 方法引入了一个协同蒸馏框架,通用模型和本地模型在框架内可以进行双向知识蒸馏,从而实现通用全局模型与多个个性化模型的协同训练. 该方法还会提取通用模型和本地模型中较低傅立叶系数的知识,由此来提高通用模型的性能. MHpFLID(Model Heterogeneous PFL via Injection and Distillation)[19]通过引入轻量级信使模型和独特的信息接收器与传输器模块,有效地解决了模型异质性和 Non-IID(Non-Independent and Identically Distributed)数据的问题. 该方法在各 FL 客户端部署信使模型,通过双向知识迁移(即本地模型吸收信使模型知识,信使模型提炼客户特征)完成本地训练,最后通过组合信使模型的参数来聚合知识. 表 2 从多个角度对比了以上几种均不依赖公共数据集的知识蒸馏方法.

表 2　5 种知识蒸馏方法对比

Table 2　Comparison of five knowledge distillation methods

方法	核心机制	优点	缺点
pFedSD[1]	本地自知识蒸馏	兼容性好	依赖历史模型质量
FedBSD[3]	基于主干网络的自蒸馏	实用性强	计算开销大
PSFL[17]	个性化本地蒸馏;自适应剪枝	平衡个性化与通信效率	计算开销大
Chen 等人[18]	傅里叶频谱相似性度量;协同蒸馏	区分共性与个性化表征	傅里叶分析较复杂
MHpFLID[19]	轻量级信使模型;双向知识迁移	针对模型异构性问题	信使模型设计复杂

2.2　基于预训练-微调

预训练通常指的是在大规模数据集上预先训练模型的过程,微调一般是指将预训练得到的模型在特定任务的有标注数据集上进行进一步训练,使其适应具体任务需求[11,15]. 在 PFL 中,微调阶段会依据各参与方的数据分布、任务需求和模型结构等对模型进行个性化调整,提升各参与方在特定任务上的性能,满足不同用户或设备的个性化需求. 图 2 为基于预训练-微调的 PFL 方法的一般流程图.

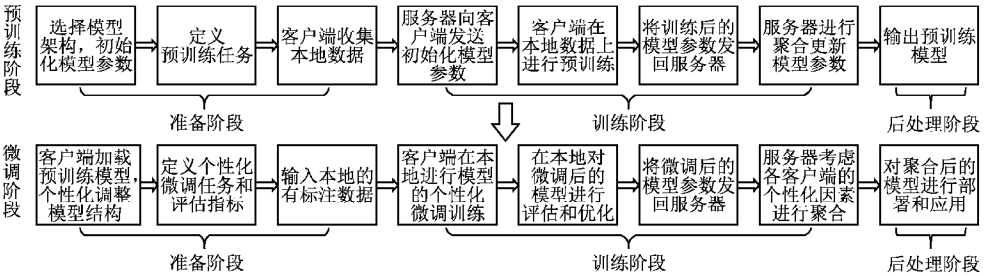

图 2　基于预训练-微调的 PFL 方法的一般流程图

Fig. 2　General flowchart of PFL methods based on pre-training and fine-tuning

预训练模型的参数数量通常非常大,简单地微调完整模型会在 FL 算法中产生巨大的通信开销[11]. 因此,很多工作以参数高效的微调方式探索大规模预训练模型的可行性,通过极少量可训练参数或冻结大部分参数,使预训练模型高效适应下游任务. FedIns[11]采用深层特征缩放与平移(scale and shift deep features,SSF)方法,为每个 FL 客户端训练了一个 SSF 池,并将它们聚合在 FL 服务器上. 在训练过程中,FL 客户端仅需传输 SSF 池中的参数,FL 服务器会动态地聚合 SSF 池的参数来生成实例专属模型,从而在达成低通信开销目标的同时实现细粒度的实例级适配. PERADA[15]在训练时充分利用预训练模型的功能,仅更新和通信适配器参数,大幅降低了计算和通信成本,并且该框架的适配器采用了一种参数高效的设计,允许轻量级的本地训练,非常适合资源受限的客户端. 为了实现较好的泛化性能,PERADA 使用全局适配器对每个 FL 客户端的个性化适配器进行正则化,全局适配器同时也负责聚合所有来自 FL 客户端的泛化信息. pFedLoRA(PFL with LoRA Tuning)[20]是一种基于低秩自适应(Low-Rank Adaptation,LoRA)模块调优的 PFL 框架,旨在解决模型

异构个性化联邦学习中的效率问题、性能问题以及灵活性需求. 每个 FL 客户端的本地异构模型中会被插入一个小的低秩同构适配器,框架运行时仅需训练适配器和传输这些适配器的参数,降低了异构模型训练的计算与通信成本. 适配器通过本地数据进行训练,不依赖公共数据集,令 pFedLoRA 的适用性更加广泛. Du 等人[21]提出的个性化联邦框架就引入了 LoRA 模块,在每一轮训练中,FL 客户端都会冻结主干模型的参数,并采用轻量级的 LoRA 模块进行参数高效的微调,以最小化通信开销. 此外,方法还结合了 Whisper 等大规模预训练模型,避免了灾难性遗忘的发生. 特别提出,FedIns 在面对跨客户端和客户端内的极端数据异构场景时准确率均较高,表现优异. 但 FedIns 和 PERADA 的性能均会受限于预训练模型的质量,若上游预训练模型的任务与联邦任务领域差异较大,方法的性能会受到影响. pFedLoRA 中的适配器为低秩线性结构,可能无法捕获复杂非线性关系,也会在一定程度上影响异构模型的个性化能力.

pFedPG(PFL via Client-Specific Prompt Generation)[4]是一种基于预训练模型的动态提示生成方法,它在训练过程中冻结预训练模型,仅需优化提示参数,同样采用参数高效的微调方式,但其方法核心在于通过 FL 服务器端的生成器学习 FL 客户端的特征,动态生成个性化提示,从而适配异构数据的分布. 在 pFedPG 框架下,每个 FL 客户端都会训练特定于其它客户端的提示,以指示模型使用其私有数据在目标 FL 客户端上执行识别任务.

参数高效微调方法的探索基础上,有研究者通过选择性更新预训练模型的敏感层,仅对模型的部分参数进行个性化本地更新调整,提出了部分模型个性化方法. FedPerfix(Federated Personalized Prefix-tuning)[5]就是其中之一,它只更新 ViT 预训练模型的自注意力层中的轻量级插件(小型的可训练参数模块),而将其余层进行全局聚合共享,利用插件来捕获 FL 客户端特定的个性化知识,在 ViT 上实现了高效的部分个性化,减少了通信与计算开销. FedPerfix 也是首次系统研究 ViT 层敏感性的方法,为 Transformer 类模型的 FL 提供了新的范式.

表 3 5 种基于预训练-微调的方法对比

Table 3 Comparison of five pre-training and fine-tuning based methods

方法	核心机制	优点	缺点
FedIns[11]	深层特征缩放与平移	通信开销极低;细粒度适配;异构数据性能好	依赖预训练模型;参数敏感
PERADA[15]	参数高效适配器	计算通信成本低;支持资源受限客户端	任务领域差异影响性能;适配器能力有限
pFedLoRA[20]	低秩自适应(LoRA)	兼容异构模型;避免灾难性遗忘;无需公共数据	线性结构限制非线性能力;需人工调参
pFedPG[4]	动态提示生成	动态适配异构数据;通信开销极低	依赖服务器特征学习;提示能力有限
FedPerfix[5]	部分模型个性化	首次系统研究 ViT 层敏感性;个性化效率高	仅适用于 Transformer;设计复杂

隐私保护迁移学习也是 PFL 的一个重要研究方向. FEDORA[22]利用自适应参数传播(仅相似客户端间传播参数)和动态选择(基于本地验证集性能决定是否接受知识迁移)的方式将 FL 重构为隐私保护的迁移学习问题,并利用分布相似性矩阵量化客户端间的关联性,减轻了 PFL 中的负迁移现象. FEDORA 还通过协变量偏移假设实现了特征空间相似性计算,使得它能够支持无标签客户端的冷启动.

另外,Fed-POE(FL with Personalized Online Ensemble)[22]令每个 FL 客户端动态融合本地模型和服务器存储的多个联邦模型,构建了一个个性化在线预测模型. 通过智能组合本地微调模型与不同阶段的全局模型,方法帮助 FL 客户端提升了模型对自身数据的适应性. 与此相对,客户端需维护本地模型 + 联邦模型 + 历史模型的子集,导致存储开销较大. 表 3 对比了 5 种具有代表性的预训练-微调方法的不同方面,展现了方法各自的优缺点和适用场景.

2.3 基于个性化特征聚合

在 FL 当中,特征聚合方法是指一种将多个 FL 客户端设备训练得到的模型特征相关信息进行整合的技术. PFL 的特征聚合方法不仅要考虑全局知识的整合,还要注重如何将全局知识与 FL 客户端的个性化特征相结合. 这里本文将个性化特征聚合方法分成两大类来区别这些聚合方法的核心思想:基于模型及参数特征的聚合方法侧重于对各参与方的模型参数进行操作和整合来得到一个全局模型,重点在于模型层面的融合与优化;基于数据相似性的聚合方法侧重于对原始数据的特征进行提取、分析和聚合,更多是从数据本身的特征角度来进行处理,以挖掘数据中的相似性和规律.

2.3.1 基于模型及参数特征聚合

FedFomo[23]首先提出了一种方案,令每个 FL 客户端下载其他 FL 客户端的模型,基于目标验证集的性能实时计算模型权重,以 FL 客户端指定任务为导向来评估其对自身目标的提升效果. 这种方法无需预知数据分布,允许每个 FL 客户端针对感兴趣的任意目标进行跨分布优化,在 Non-IID 场景下实现了高效的个性化,突破了传统方法对数据分布一致性的依赖. 虽然 FedFomo 方法的性能不错,但框架中的 FL 客户端需下载多个模型,导致了高通信开销和隐私风险;同时 FL 客户端需要维护验证集并计算模型性能,这也增加了本地计算的负担. 为了在不增加额外通信开销和隐私泄露风险的情况下准确获取每个 FL 客户端下载的全局模型中的信息,FedALA(FL with Adaptive Local Aggregation)[12]让 FL 客户端从 FL 服务器端下载全局模型,通过自适应聚合模块将全局模型与旧的本地模型进行本地聚合来进行初始化,并将训练后的本地模型上传到 FL 服务器端. 值得注意的是,FedALA 首次提出了元素级权重的概念,与 FedFomo 使用模型级权重进行粗粒度聚合相对,实现了对于参数的细粒度控制,进一步提升了模型的个性化性能.

上述两种方法选择聚合全局和本地模型的全部参数,这样做可能会引入噪声. 因此,FedAFK(PFL with Adaptive Feature Aggregation and Knowledge Transfer)[24]选择仅聚合特征提取器,保留分类头的本地化特性,通过可学习的权重系数来提升个性化的灵活性,减少通信开销. 在本地特征提取器的训练目标中,FedAFK 还加入了一个附加项以传递全局特征表示中包含的知识,从而在特征层面平衡了全局泛化与本地个

性化. 不过,该方法需同时训练全局特征提取器、本地特征提取器、分类头和自适应权重参数,单轮次的计算时间会高于部分 PFL 方法.

在创新层面,pFedGPA[25]突破了传统基于调整或选择的个性化范式,旨在通过扩散模型实现参数空间的生成式聚合. 方法在 FL 服务器上部署了一个扩散模型来聚合不同的参数分布,并提出了一种参数反演方法来为每个 FL 客户端生成一组个性化参数. 尽管扩散模型训练较为耗时,在大规模客户端场景可能成为系统瓶颈,pFedGPA 依然革新了 FL 的参数聚合范式,为 FL 与生成式 AI 的交叉提供了新思路. FedC^{2}I[26]则创新了一种动态聚合机制. 在全局共享层面,特征表示层通过 FL 客户端级的影响向量实现了动态聚合;在本地个性化层面,分类器通过类级影响矩阵实现了自适应融合. FedC^{2}I 中的影响度权重计算由 FL 客户端自主完成,无需 FL 服务器端干预,使得该方法更加适合隐私敏感场景. FedSelect[27]另辟蹊径,从彩票假设(Lottery Ticket Hypothesis, LTH)中汲取灵感,提出一种新颖的基于梯度变化的参数重要性评估方式,对在训练中变化幅度大的参数进行个性化,对变化幅度小的参数进行全局聚合和共享. 不同于传统的固定层级参数解耦,FedSelect 通过动态选择参数子集进行个性化,摆脱了对于人工预设层结构环节的依赖,为 Non-IID 数据下的 FL 提供了新思路. FedCP(Federated Conditional Policy)[28]跳出依赖模型参数传递信息的思维定式,直接操作数据特征,使用条件策略网络(Conditional Policy Network,CPN)从源头显式分离特征中的全局与个性化信息,旨在实现样本级别的特征分配,令方法能够适应更复杂的异构场景,为 PFL 提供了新的范式. 同样是针对特征中的全局与个性化信息,GPFL[29]打破了二者的对立,通过条件阀(Conditional Valve, CoV)分离了两种特征信息的提取路径,并引入全局类别嵌入层(Global Category Embedding layer,GCE),从幅度和角度两个维度来引导特征对齐,达成了同步高效提取、优化全局和个性化特征信息的目标,在协作与个性化间取得了突破性的平衡,有效地抑制了个性化模型的过拟合,增强了模型的公平性和隐私保护能力.

2.3.2　基于数据相似性聚合

FedAMP(Federated Attentive Message Passing)[30]致力于促进具有相似数据及数据分布的 FL 客户端之间进行成对协作,方法的核心在于联邦注意力消息传递机制. FedAMP 在云端 FL 服务器上为每个客户端维护一个个性化模型,并智能地筛选出数据及数据分布相似的客户端模型,在这些成对的模型之间进行定向信息传递,从而实现联邦注意力消息传递机制. FedAMP 还采用相似度加权的动态融合策略持续优化模型,这让 FL 客户端能够自动寻找协作学习的最佳伙伴,形成了自组织的高效协作网络,实现了去中心化协作的目标. FedAMP 在实验中展示了良好的性能,但实验集中于静态异构数据,方法并未对协作范围进行动态调整,这可能会导致引入噪声或错失潜在的协作机会. 为了解决这个问题,FedCAC[31]将对 Non-IID 敏感的参数确定为关键参数,允许关键参数在早期与更多 FL 客户端进行协作,并逐步将协作对象缩小至相似的 FL 客户端,通过时间阈值机制动态调整关键参数的协作范围,从而适配不同训练阶段的需求. 通过对参数敏感性和分布相似性的双重建模,FedCAC 实现了更高效的跨客户端协作,且在复杂 Non-IID 场景下表现尤其突出. FedReMa[13]同样考虑了随时间改变协作策略的可能性,提出了关键协作周期的概念,阶段化调整协作和聚合策略,达成了个性化和泛化之间的平衡. 方法利用 FL 客户端分类器的 logits(分类模型最后一层的原始输出值)相似性来识别任务之间的相关性,在动态筛选高相关 FL 客户端的同时降低了计算开销,适应了类不平衡场景.

上述几种方法都仅关注了数据分布相似的客户端,但忽视了不同分布客户端的潜在价值,这可能导致模型陷入局部最优. DiversiFed[32]提出了一种更具有创新性的协作方案,其核心是通过一个模型距离损失函数使相似分布的客户端模型在参数空间中靠近,共享局部知识,同时让不同分布的模型在参数空间中被推离,避免他们之间的冲突并捕捉有价值的互补信息. 通过这种动态平衡的推拉策略,DiversiFed 利用不同分布的 FL 客户端的信息实现了多样性增益,大大提高了方法在高度 Non-IID 场景下的性能.

2.4　基于模型正则化

在 FL 当中,模型正则化方法可以通过限制模型参数的大小来控制模型复杂度,防止模型出现过拟合现象,从而提高模型的泛化能力、稳定性和收敛速度. 将模型正则化方法运用到 PFL 上,需要仔细考虑如何平衡全局模型和 FL 客户端本地模型的差异.

Ditto[33]在本地目标中增加正则项,通过正则化手段将本地模型向全局模型对齐,对个性化和泛化进行了有效的平衡. 在解耦优化过程中,框架内的全局模型和本地模型交替更新,取代了传统的参数耦合方法,降低了过程复杂度,在 PFL 中实现了公平性、鲁棒性与准确性的协同提升. 遗憾的是,方法并不能够很好地适应异构环境下的 PFL.

针对 Ditto 方法的痛点,pFedMe(PFL with Moreau Envelopes)[34]将莫罗包络(Moreau envelope)作为客户的正则化损失函数,将优化个性化模型的过程与学习全局模型的过程解耦,降低了优化个性化模型时的复杂度,并利用莫罗包络的保凸性和平滑性来促进 pFedMe 的收敛性分析,有效地提高了算法的收敛速度和准确性,解决了统计异构性下的 PFL 问题.

向模型中添加正则项可以使模型参数稀疏化,由此构建出稀疏模型. pFedGate[35]为每个客户端引入了一个轻量级的可训练门控层,利用轻量级门控层动态生成适配不同 FL 客户端数据分布的稀疏模型,通过模型的稀疏化降低计算和通信开销;同时允许 FL 客户端根据自身资源量动态调整稀疏模型容量,突破了低资源 FL 客户端的容量限制,解决了 PFL 面对资源异构性时的难点. pFedGate 的动态适配能力使其在跨设备联邦场景中具有一定的应用潜力.

2.5　基于元学习

元学习是一种能够快速适应新任务或新数据分布的学习策略,通常通过构建元模型来实现. 元模型会在训练过程中学习如何根据不同 FL 客户端的数据特征和任务特性来调整模型的参数. 在 PFL 中,元学习可以充分利用每个 FL 客户端的数据特点和任务需求来训练个性化的模型,提升模型在 FL 客户端上的个性化性能.

传统的模型不可知的元学习(Model-Agnostic Meta-

Learning,MAML)方法通常假设任务来自同一分布且数据集中存储,而 Per-FedAvg(Personalized FedAvg)[36]将其扩展到了 FL 的分布式场景,量化了数据分布差异对收敛的影响,揭示了梯度估计偏差与方差的关键性质,为后续的研究提供了理论基石. Per-FedAvg 的核心在于将 MAML 与联邦架构结合,通过元学习框架使全局模型能够作为 FL 客户端本地模型的初始化点,用户仅需少量本地训练即可快速适配自身数据,发挥出了 FL 在隐私与通信效率方面的优势.

尽管 Per-FedAvg 创新十足,但其在异构数据下仍然存在偏差积累问题. PFLDyn[37]通过动态调整本地损失函数和梯度校正的方式,减少了设备数据偏差对全局元模型的影响,解决了全局元模型与本地个性化目标的对齐问题,显著提升了通信效率和个性化性能. 更重要的是,面对设备数据类别高度不重叠和标签匿名等极端异构环境时,PFLDyn 仍然可以保持高性能. 上述两种方法都是元学习、FL 和 PFL 的结合成果,都拥有良好的性能,但对设备算力要求较高,仍然有改进空间.

2.6 混合专家模型

混合专家模型(Mixture of Experts,MoE)是一种深度学习架构,它由多个专家子模型和一个门控网络组成,每个门控网络和专家池都是一个前馈网络[14]. 所有专家子模型都有独立的参数和结构,专注于学习数据的某一个特定方面或者某一个子任务,而门控网络通常是一个简单的神经网络,负责接收输入数据并输出每个专家的重要性权重. 在 PFL 中,MoE 可以让不同的专家学习不同 FL 客户端数据的特征,为每个 FL 客户端构建个性化的专家组合,满足个性化的需求. 图 3 展现了基于 MoE 的 PFL 方法的一般架构.

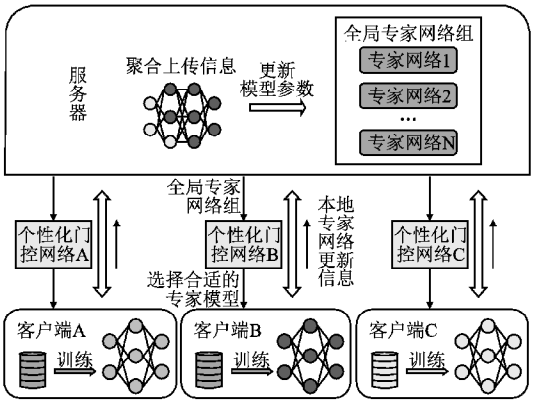

图 3 基于 MoE 的 PFL 方法的一般架构图

Fig. 3 General architecture diagram of PFL methods based on MoE

现阶段已经有许多研究将混合专家模型与 PFL 结合在一起. FedMoE[14]用混合专家模型取代了传统的密集模型,通过 MoE 的稀疏激活机制和一种动态子模型调整方法,为每个 FL 客户端动态构建了异构的子模型. 这种调整方法分为启发式搜索和专家推荐两个阶段,用于实现子模型从“次优”到“最优”的渐进式动态优化. FedMoE 还把专家模块作为基本单元,允许更细粒度个性化的同时也支持进行模块级聚合,仅聚合激活的专家参数,显著降低了通信和内存消耗,非常适合跨任务的异构场景.

FedMoE 在多方面表现出色,但其分为两阶段的动态子模型调整方法仍然略显复杂,需要花费额外的计算资源来执行启发式搜索等步骤. FedMix[38]基于输入或标签信息,通过 MoE 自适应地动态分配 FL 客户端到特定专家处,再结合变分推断来优化专家选择,实现了能够同时处理多种 Non-IID 场景的目标,在性能相差不大的前提下使整个方法的流程相比 FedMoE 复杂度更低,计算开销更小. 另外,FedMix 令具有相似数据的 FL 客户端自动选择相同的专家,降低了方法对于数据标签的依赖性.

FedMix 方法需要人工指定专家数量 K,且在训练过程中需传输 K 倍的模型参数,通信开销较高,不适合跨设备场景. Zadouri 等人[39]提出了一种参数高效的 MoE 架构,将混合专家模型架构与轻量级专家相结合,使方法在极低的参数预算(<1%)下达到甚至超越了与全微调相当的性能,为大规模模型的高效部署提供了新思路. 架构中的软路由机制可根据 FL 客户端的本地数据动态选择专家组合,从而适配不同 FL 客户端的个性化需求,提升个性化性能.

表 4 4 种 MoE 架构方法的对比

Table 4 Comparison of four MoE architecture methods

方法	核心架构	专家功能	门控网络功能	训练方式
Fed-MoE[14]	稀疏激活 MoE	学习客户端特征,构建个性化专家组合	识别次优子模型,生成个性化权重	启发式搜索+模块化聚合训练
MoLO-RA[39]	MoE+轻量级专家	参数高效更新,解决规模挑战	在路由器中按门控分数组合专家输出,实现条件计算	模块化参数更新,极低参数量微调
Fed-Mix[38]	MoE+门控机制	自适应选择并训练用户特定模型成员	按数据子集调节模型行为,实现个性化	训练专家模型,客户端识别相关专家
pFed-MoE[40]	MoE+本地异构模型+共享特征提取器+本地门控网络	本地异构特征提取器为个性化专家,共享同构特征提取器为全局专家	为每个样本生成个性化权重,融合两位专家特征	本地训练共享同构特征提取器并上传服务器进行聚合

pFedMoE(PFL with Mixture of Experts)[40]利用 MoE 来针对性地解决模型异构性的问题,将 FL 客户端中的小型同构特征提取器作为全局专家进行共享,将本地异构特征提取器作为局部专家. 在训练时,FL 客户端中的门控网络为每个样本生成个性化权重,动态融合全局特征与局部特征,实现了在保护模型隐私的同时高效融合知识的目标,提升了方法的个性化性能. pFedMoE 为模型异构场景下的 PFL 提供了新的思路,在资源受限的边缘计算中具有很大的应用潜力. 表 4 对上述 4 种 MoE 架构的方法进行了多方面的对比.

2.7 其他类别

除了上述几类,还有很多研究站在较为独特的角度提出了一些有价值的 PFL 方法,这里将它们统一放在其他类别当中进行介绍.

1)双掩码机制. 现有的 PFL 解决方案很容易受到训练和测试数据之间分布偏移的影响. DM-PFL(Dual Masked PFL)[41]设计了一种新颖的双掩码机制,利用该机制实现了权重级参数共享和端到端动态稀疏训练,为每个 FL 客户端训练了一个全局模型和一组个性化模型. 方法通过共享稀疏

参数减少冗余计算,以极小的额外训练开销提高了模型的迁移鲁棒性,降低了数据分布偏移对于模型个性化性能的影响.由于较高的稀疏率可能导致模型出现欠拟合状况,方法仍需进行优化来更好地平衡稀疏度与个性化性能.

2)*抵御后门攻击*. Simple-Tuning[42]对 PFL 框架中的后门攻击进行了研究,证明具有部分模型共享功能的 PFL 方法可以显著提高对后门攻击的鲁棒性,揭示了模型共享程度与防御能力的直接关联,由此提出了一种可以有效抵御后门攻击的轻量级防御方法,为轻量防御设计提供了新思路. Simple-Tuning 只对线性分类器进行调整,对每个 FL 客户端的训练模型进行线性分类,因此更容易与现有的 FL 方法结合,降低了计算成本.但对于部分大模型,Simple-Tuning 的防御效果不甚理想,需要结合其他手段进行协同防御.

3)*相互学习和集成学习*. FedAPEN(Federated Adaptive Personalized Ensemble)[43]将 PFL 与相互学习和集成学习结合在一起,有效利用了私有模型和共享全局模型的优势.方法创新地提出了两种自适应集成机制:适应性学习机制允许每个 FL 客户端单独学习一个权重系数,模型在集成过程中利用这个系数进行自适应加权;集成学习机制用于提高模型集成的预测准确性,实现了 FedAPEN 对于统计和模型异构性的双重适应.不过,FedAPEN 在运行过程中需要同时维护私有模型和共享模型,这增加了它的训练时间和内存消耗.

4)*个性化语义融合*.现有的 PFL 方法在进行局部微调时容易将全局模型中的判别性语义覆盖掉,导致模型的个性化性能下降. Xia 等人[44]提出了一种个性化语义激励机制,通过梯度差异来筛选出全局模型中未被充分激活的通道,并动态调整局部模型的参数,在语义层面对全局知识进行选择性融合,从而保护全局判别知识.方法中的跨模型注意力模块利用全局模型的通道相似性,有效提升了分类器对于数据分布偏移的适应性,在个性化与全局知识保留之间取得了良好的平衡效果.本文方法成立的前提是全局模型在本地训练期间固定,但在实际的 FL 场景中全局模型可能会进行动态更新,因此该理论还需进一步发展.

5)*基于逻辑属性聚类*.当前的 FL 领域缺少对 FL 客户端模型的符号推理能力的研究. FedSTL(Signal Temporal Logic-enabled PFL)[45]为存在分歧的 FL 客户端自动推断时序逻辑属性,增强模型的时间推理能力,并依据时序逻辑属性的相似性对 FL 客户端进行聚类,实现符号推理与数据驱动的 FL 深度融合,提升预测准确性和逻辑一致性.此外,FedSTL 增强了特定的数据聚类模式和 FL 客户端属性配置,自动从数据中提取 FL 客户端及集群的时间和逻辑特征,在实现多维度个性化定制的同时也显著降低了 FL 客户端固有异构性对系统性能的影响.与此相对,FedSTL 中的属性推断和动态聚类环节增加了每轮通信的计算负载,对于资源受限的设备不够友好,因此之后需要优化方法中的属性挖掘算法,降低计算开销.

3　新兴方向

3.1　个性化联邦长尾学习

长尾分布是指数据集的一种分布情况,少量类别占据了大部分数据样本,而大量类别仅有少量数据样本,长尾学习就是针对呈现长尾分布数据的学习方法.当 FL 客户端的数据呈现长尾分布时,就需要用到联邦长尾学习. FedLoGe(Federated Local and Generic Model Training)[46]探索了这一方向,通过共享全局特征提取器和本地个性化分类器的解耦设计,在数据呈现长尾分布的情形下同时提升了全局通用模型和本地个性化模型的性能.方法创新地引入神经崩溃框架,利用固定分类器的几何约束来引导特征空间的优化,为后续研究提供了新思路. SSE-C(Static Sparse Equiangular Tight Frame Classifier)模块使用稀疏剪枝策略实现了噪声特征过滤与高效的表示学习,GLA-FR(Global and Local Adaptive Feature Realignment)模块则分别针对全局和本地进行特征-分类器的对齐,减少了异构客户端的特征分歧.当然,FedLoGe 也有不足之处,SSE-C 中稀疏策略的自适应改变、保护全局分类器范数中隐含的本地数据隐私等都是其后续可以优化的方面.

3.2　个性化联邦子图学习

图联邦学习是 FL 和图数据挖掘相结合的一种新兴技术,是指在保护数据隐私的前提下利用分布式的图数据进行联邦学习. FED-PUB(Federated Personalized Subgraph Learning)[47]从个性化子图联邦学习角度出发,利用功能嵌入来计算模型相似性,动态划分出具有相似子图的 FL 客户端社区,实现社区内的知识共享.方法还利用 FL 客户端本地的稀疏掩码,筛选全局聚合参数中与本地子图相关的子网络,从而实现参数级个性化,解决了子图联邦学习中的隐私-效率权衡问题.然而在极端异构场景中,FED-PUB 的相似性度量容易失效,性能下降明显. FedSheafHN(FL with Sheaf Diffusion and HyperNetworks)[48]通过 3 个创新点解决了这个问题.首先,方法利用 FL 客户端子图的图级嵌入进行协作图建模,将 PFL 问题转化为图学习任务.其次,FedSheafHN 通过微分几何中的鞘理论对 FL 客户端间的非线性关系进行建模,显式捕捉 FL 客户端间的隐式拓扑关系.最后,方法结合注意力增强的超网络实现了跨 FL 客户端信息的动态聚合. FedSheafHN 在医疗协作网络、跨平台推荐系统等需要保护隐私、数据分布高度差异化的场景中有显著的应用潜力.

3.3　个性化联邦域增量学习

域增量学习主要关注在不同的数据域之间进行增量学习,模型会不断地从新的数据样本中学习知识,逐步更新自己的参数以适应新的信息.在 FL 中,各个数据域也会经常发生增量变化,联邦域增量学习应运而生. pFedDIL(Personalized Federated Domain-Incremental Learning)[49]作为个性化联邦域增量学习方法,旨在解决学习过程中的灾难性遗忘问题.方法提出自适应知识匹配机制,利用辅助分类器量化了新任务与历史任务的相似性.在训练新任务时,FL 客户端可以根据相似性来动态选择采用新的初始模型或者对旧模型进行复用,从以前的任务中迁移知识,大大减少了不必要的训练量,在保证隐私的前提下实现了跨领域任务的高效增量学习.此外,pFedDIL 中的目标分类模型与辅助分类器会共享底层的特征提取层,以便于压缩模型参数量,降低系统开销. pFedDIL 的性能依赖于任务相关性假设,因此若新任务与历史任务差异极大,其知识迁移方法可能会失效,性能也会下降.

3.4　针对时空异质性的个性化联邦学习

时空异质性是指数据在空间和时间维度上呈现出的非均

匀特性,这种特性使得各 FL 客户端的数据分布和变化规律存在差异,会导致模型收敛困难、性能下降. PFL 可以根据不同 FL 客户端的时空特性,为每个 FL 客户端定制个性化的模型训练和更新策略,是解决该问题的有效方法. FUELS(Federated Dual Semantic Alignment-based Contrastive Learning)[50] 作为第一个针对时空异质性的 PFL 方法,核心创新点在于动态语义对齐与轻量原型通信. 方法通过设计时间或空间维度的对比任务,利用对比学习对时空异质性进行了直接建模,同时在表示空间中对齐了语义相似样本,分离了不相似样本,有效增强了 FL 客户端模型的个性化表达能力. FUELS 还用能够感知流量周期性的原型(prototype)替代全模型参数作为通信载体,压缩了 FL 客户端的特征信息和传输的数据量,极大降低了通信成本,为边缘计算环境下的时空预测任务提供了新的解决方案.

3.5 个性化联邦持续学习

持续学习是指机器学习系统不断从新数据中学习知识,且在学习新知识时不会忘记之前学到的旧知识. 与增量学习不同,持续学习需要重点考虑知识的遗忘问题,而增量学习相对更关注如何快速有效地利用新数据更新模型. FedMGP(Federated Multi-Granularity Prompt)[51] 是一种基于多粒度知识表示的个性化联邦持续学习方法,把预训练 ViT 大模型作为共享认知的基础,通过输入层的全局提示和自注意力层的本地提示来分离共享知识和个性化知识. 方法利用全局提示来保留时空不变的知识,利用本地提示来捕获任务特定的知识,避免学习出现时空灾难性遗忘情况. 在知识融合过程中,FedMGP 通过一种选择性提示融合机制,仅聚合各 FL 客户端的粗粒度全局提示知识,避免了本地细粒度知识的干扰,在联邦持续学习场景中首次实现了时空知识的高效融合与个性化适配. 不过,FedMGP 的本地 FL 客户端需要维护一个全局 + 本地的提示池,参数量极大,后续可以通过低秩分解等提示压缩技术来压缩参数量,降低训练复杂度.

3.6 个性化联邦多模态学习

多模态学习通过处理和整合来自多种模态的数据来完成任务,联邦多模态学习则是近几年的新兴研究方向,旨在利用多个 FL 客户端设备上的多模态数据进行协作学习. Yin 等人[52] 对联邦多模态学习的模型聚合过程进行了改进,提出基于学习的聚合系数优化方案,根据设备的数据相似性和模态特性动态调整参数权重,优先聚合数据相似设备的参数. 方法还针对模态异构性和 Non-IID 数据设计了一种模态级参数调度策略,仅上传关键的模态参数,减少了参数传输量. 在通信开销方面,方法将参数上传的决策与信道状态结合,兼顾了个性化性能与通信效率,为个性化联邦多模态学习领域打开了新的思路.

3.7 个性化量子联邦学习

量子学习是将量子力学原理和计算方法应用于学习过程的一种新兴技术,利用量子态的叠加、纠缠等特性来处理和存储信息. 量子联邦学习以 FL 为基础,在数据加密和模型聚合等阶段利用量子计算的优势来提升 FL 的性能. Shi 等人[53] 提出了一种个性化量子联邦学习方法,其核心是修改模型结构,引入个性化层让 FL 客户端模型保留本地特征,从而提高模型的个性化性能. 此外,方法设计了一个安全协议,利用量子不可克隆定理来抵御内部和外部的攻击,实现对模型参数和数据的双重隐私保护,为分布式量子机器学习提供了新思路. 方案需要 FL 客户端具备量子计算能力,但目前的 NISQ(Noisy Intermediate-Scale Quantum,含噪声中等规模量子)设备难以支持多 FL 客户端并行训练,未来可以探索在 FL 客户端使用量子模拟器来降低部署门槛的方法.

4 领域应用

在理论方法与技术框架不断创新的同时,PFL 的实用价值也在诸多对数据隐私与个性化需求并重的垂直领域中得到凸显,为打破数据孤岛、在保护隐私的前提下实现协同智能提供了可行的技术路径. 本文接下来将分别探讨 PFL 在医学、语音识别、金融及气象领域的典型应用场景与解决方案.

4.1 医学领域

医学领域会涉及到大量患者的敏感信息,因此一直对隐私保护非常重视. FL 可以在不泄露患者隐私的前提下进行数据协作和模型训练,更容易获得患者的信任和认可. 前文在基于知识蒸馏分类中曾提到一种模型异构的个性化注入蒸馏联邦学习方法 MHpFLID[19],研究尝试将其应用在医学领域,以消除模型训练对公共医疗数据集的依赖以及本地训练相关的额外成本,减轻医疗机构在本地计算和存储资源方面的负担. MHpFLID 在医学图像分类、医学图像分割和医学时间序列分类等多个医学任务上进行的验证结果都十分理想,在医学领域具有较大的应用潜力.

糖尿病问题在医学界备受关注,定期监测血糖值是改善和预防糖尿病的重要手段,其中识别低血糖和高血糖等罕见事件具有一定挑战性,对敏感患者数据的获取受限也阻碍了稳健的机器学习模型的开发. 为了解决这些问题,Dave 等人[54] 提出了一种新的 Hypo-Hyper(HH)损失函数和一个 PFL 框架 FedGlu. Hypo-Hyper 损失函数的运用显著提高了血糖预测的准确性,降低了血糖偏移预测问题带来的危害. FedGlu 框架允许 FL 客户端在本地训练模型,并在其他患者之间仅共享模型参数来进行协作学习,无需共享敏感数据,体现了框架的应用潜力.

4.2 语音识别领域

语音中包含了大量的敏感信息,常常涉及到个人或者团体的重要隐私,且语音数据分散在众多用户的设备上,形成了数据孤岛. PFL 可以在保护用户数据隐私的同时根据不同用户的语音使用习惯和需求来定制个性化的语音识别模型,提高识别准确率. 现有的联邦语音转文本(Speech-to-Text,S2T)工作流涉及 FL 服务器和 FL 客户端之间的多轮整体模型交互,这种方式在产生大量计算和通信开销的同时也令 FL 客户端在异构环境下难以适应整个模型训练和通信的严格要求. Du 等人[21] 通过引入 LORA 模块来减少通信开销,针对 FL 客户端的记忆检索机制来实现个性化. 这种方式能够在大部分 S2T 任务中显著降低通信开销,并有效地对全局模型进行个性化处理,很大程度上克服了数据异质性问题.

4.3 金融领域

金融领域的数据敏感性非常高,涉及到大量客户隐私和商业机密,且这些数据通常具有高度异质性,不同机构或业务

场景的数据分布和特征差异大. 在金融领域,PFL 能够在 FL 保证数据安全的基础上根据各方特点进行定制化模型训练. YU 等人[55]提出了 PFL 方法 Fed + . 该方法不要求所有参与方都收敛到同一个中心点,允许中央服务器采用稳健的方法聚合本地模型,同时保持本地计算结构不变,从而使方法能更好地适应现实世界中数据的各种特性. 在金融投资组合管理中,Fed + 让投资组合经理在不泄露私有数据下共同训练策略,以隐私保护方式实现多任务学习,提高了模型在金融市场非平稳性下的可迁移性,解决了标准 FL 方法因参与方数据异质性导致的训练失败问题.

4.4　气象领域

气象部门的观测站数据可能涉及国家关键基础设施周边的气象细节,这些数据的隐私和安全至关重要,且气象数据广泛分布在世界各地的众多由不同的机构或部门管理的数据源上,这些特点都决定了 FL 在气象领域大有可为. adapFL(adaptive FL)[56]是一种适用于分布式气象雷达图像的 PFL 框架. 该框架假设有一台覆盖直径达 400 公里的可用雷达,雷达捕获的图像被划分为 4 个象限区域,以此模拟每部雷达覆盖整个区域中的一个子区域的场景,这些区域会互不重叠地分配到 4 个不同的 FL 客户端中. 每个区域以及覆盖先前分配的各个区域部分表面的中心区域都会对通过 adapFL 所获得的结果进行分析.

5　总　结

随着信息技术的飞速发展,数据的隐私保护与有效利用成为了各界关注的焦点,个性化联邦学习应运而生,为这一难题提供了创新性的解决方案. 通过本文对其的全面综述可知,现有的 PFL 方法主要聚焦于如何在保护数据隐私的 FL 框架下有效实现个性化定制,使得各参与方既能受益于全局知识共享,又能贴合自身数据特性来优化模型.

数据异构性、模型融合问题以及个性化与全局模型的平衡等成为了制约个性化联邦学习发展的重要挑战. 为应对这些难题,一系列 PFL 方法蓬勃发展,从基于知识蒸馏、预训练-微调、个性化特征聚合等不同角度出发,各有优劣,在不同应用场景和数据集下展现出各异的性能表现,为解决实际问题提供了多样化的思路.

References:

[1] Jin H, Bai D, Yao D, et al. Personalized edge intelligence via federated self-knowledge distillation[J]. IEEE Transactions on Parallel and Distributed Systems, 2022, 34(2): 567-580.

[2] Wu J, Bao W, Ainsworth E, et al. Personalized federated learning with parameter propagation[C]//Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2023: 2594-2605.

[3] Wang P, Liu B, Zeng D, et al. Personalized federated learning via backbone self-distillation[C]//Proceedings of the 5th ACM International Conference on Multimedia in Asia (MMAsia), 2023: 1-7.

[4] Yang F E, Wang C Y, Wang Y C F. Efficient model personalization in federated learning via client-specific prompt generation[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2023: 19159-19168.

[5] Sun G, Mendieta M, Luo J, et al. Fedperfix: towards partial model personalization of vision transformers in federated learning[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2023: 4988-4998.

[6] Mclaughlin C, Su L. Personalized federated learning via feature distribution adaptation[C]//Advances in Neural Information Processing Systems (NeurIPS), 2024: 77038-77059.

[7] Liu Y, Wu X, Liu Chengkun. Research on personalized federated learning algorithm in industrial internet of things[J]. Journal of Chinese Computer Systems, 2025, 46(1): 209-216.

[8] Huang Yuchen, Zhao Yanchao, Hao Jiangshan, et al. Research on performance optimization of federated learning for data heterogeneity[J]. Journal of Chinese Computer Systems, 2024, 45(4): 777-783.

[9] Ye M, Fang X, Du B, et al. Heterogeneous federated learning: state-of-the-art and research challenges[J]. ACM Computing Surveys, 2023, 56(3): 1-44.

[10] Guo Guijuan, Tian Hui, Pi Huijuan, et al. Advances in federated learning for non-independent identically distributed data[J]. Journal of Chinese Computer Systems, 2023, 44(11): 2442-2449.

[11] Feng C M, Yu K, Liu N, et al. Towards instance-adaptive inference for federated learning[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2023: 23287-23296.

[12] Zhang J, Hua Y, Wang H, et al. Fedala: adaptive local aggregation for personalized federated learning[C]//Proceedings of the AAAI Conference on Artificial Intelligence (AAAI), 2023: 11237-11244.

[13] Liang H, Zhan Z, Liu W, et al. FedReMa: improving personalized federated learning via leveraging the most relevant clients[M]. Netherlands: IOS Press, 2024.

[14] Mei H, Cai D, Zhou A, et al. FedMoE: personalized federated learning via heterogeneous mixture of experts[J]. arXiv preprint arXiv: 2408.11304, 2024.

[15] Xie C, Huang D A, Chu W, et al. Perada: parameter-efficient federated learning personalization with generalization guarantees[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024: 23838-23848.

[16] Chen Cong, Li Jing. Knowledge distillation for federated learning with Non-IID data[J]. Journal of Chinese Computer Systems, 2025, 46(6): 1289-1297.

[17] Peng Y, Jiang F, Dong L, et al. Personalized federated learning for generative AI-assisted semantic communications[J]. arXiv preprint arXiv: 2410.02450, 2024.

[18] Chen Z, Yang H, Quek T, et al. Spectral co-distillation for personalized federated learning[C]//Advances in Neural Information Processing Systems (NeurIPS), 2023: 8757-8773.

[19] Xie L, Lin M, Luan T, et al. MH-pFLID: model heterogeneous personalized federated learning via injection and distillation for medical data analysis[C]//Proceedings of the 41st International Conference on Machine Learning (ICML), 2024: 54561-54575.

[20] Yi L, Yu H, Wang G, et al. pFedLoRA: model-heterogeneous personalized federated learning with LoRA tuning[J]. arXiv preprint arXiv: 2310.13283, 2023.

[21] Du Y, Zhang Z, Yue L, et al. Communication-efficient personalized federated learning for speech-to-text tasks[C]//IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2024: 10001-10005.

[22] M Ghari P, Shen Y. Personalized federated learning with mixture of models for adaptive prediction and model fine-tuning[C]//Ad-

vances in Neural Information Processing Systems,2024:92155-92183.
[23] Zhang M,Sapra K,Fidler S,et al. Personalized federated learning with first order model optimization[C]//International Conference on Learning Representations(ICLR),2021:10622-10638.
[24] Yin K,Mao J. Personalized federated learning with adaptive feature aggregation and knowledge transfer[J]. arXiv preprint arXiv:2410.15073,2024.
[25] Lai J,Li J,Xu J,et al. pFedGPA:diffusion-based generative parameter aggregation for personalized federated learning[C]//Proceedings of the AAAI Conference on Artificial Intelligence(AAAI),2025:17999-18007.
[26] Tan Y,Long G,Jiang J,et al. Influence-oriented personalized federated learning[J]. arXiv preprint arXiv:2410.03315,2024.
[27] Tamirisa R,Xie C,Bao W,et al. Fedselect:personalized federated learning with customized selection of parameters for fine-tuning[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR),2024:23985-23994.
[28] Zhang J,Hua Y,Wang H,et al. Fedcp:separating feature information for personalized federated learning via conditional policy[C]//Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining,2023:3249-3261.
[29] Zhang J,Hua Y,Wang H,et al. Gpfl:simultaneously learning global and personalized feature information for personalized federated learning[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision(ICCV),2023:5041-5051.
[30] Huang Y,Chu L,Zhou Z,et al. Personalized cross-silo federated learning on non-IID data[C]//Proceedings of the AAAI Conference on Artificial Intelligence(AAAI),2021:7865-7873.
[31] Wu X,Liu X,Niu J,et al. Bold but cautious:unlocking the potential of personalized federated learning through cautiously aggressive collaboration[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision(ICCV),2023:19375-19384.
[32] Wu X,Liu X,Niu J,et al. The diversity bonus:learning from dissimilar distributed clients in personalized federated learning[J]. arXiv preprint arXiv:2407.15464,2024.
[33] Li T,Hu S,Beirami A,et al. Ditto:fair and robust federated learning through personalization[C]//International Conference on Machine Learning(ICML),2021:6357-6368.
[34] T Dinh C,Tran N,Nguyen J. Personalized federated learning with moreau envelopes[C]//Advances in Neural Information Processing Systems(NeurIPS),2020:21394-21405.
[35] Chen D,Yao L,Gao D,et al. Efficient personalized federated learning via sparse model-adaptation[C]//International Conference on Machine Learning(ICML),2023:5234-5256.
[36] Fallah A,Mokhtari A,Ozdaglar A. Personalized federated learning:a meta-learning approach[C]//Proceedings of the 34th Conference on Neural Information Processing Systems(NeurIPS),2020,doi:10.48550/arXiv.2002.07948.
[37] Acar D A E,Zhao Y,Zhu R,et al. Debiasing model updates for improving personalized federated training[C]//International Conference on Machine Learning(ICML),2021:21-31.
[38] Reisser M,Louizos C,Gavves E,et al. Federated mixture of experts[J]. arXiv preprint arXiv:2107.06724,2021.
[39] Zadouri T,Üstün A,Ahmadian A,et al. Pushing mixture of experts to the limit:extremely parameter efficient MoE for instruction tuning[C]//12th International Conference on Learning Representations(ICLR),2024:22214-22233.
[40] Yi L,Yu H,Ren C,et al. pFedMoE:data-level personalization with mixture of experts for model-heterogeneous personalized federated learning[J]. arXiv preprint arXiv:2402.01350,2024.
[41] Zhang W,Zhou Z,Wang Y,et al. Dm-pfl:hitchhiking generic federated learning for efficient shift-robust personalization[C]//Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining,2023:3396-3408.
[42] Qin Z,Yao L,Chen D,et al. Revisiting personalized federated learning:robustness against backdoor attacks[C]//Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining,2023:4743-4755.
[43] Qin Z,Deng S,Zhao M,et al. Fedapen:personalized cross-silo federated learning with adaptability to statistical heterogeneity[C]//Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining,2023:1954-1964.
[44] Xia H,Li K,Ding Z. Personalized semantics excitation for federated image classification[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision(ICCV),2023:19301-19310.
[45] An Z,Johnson T T,Ma M. Formal logic enabled personalized federated learning through property inference[C]//Proceedings of the AAAI Conference on Artificial Intelligence(AAAI),2024:10882-10890.
[46] Xiao Z,Chen Z,Liu L,et al. FedLoGe:joint local and generic federated learning under long-tailed data[C]//12th International Conference on Learning Representations(ICLR),2024:32383-32402.
[47] Baek J,Jeong W,Jin J,et al. Personalized subgraph federated learning[C]//Proceedings of the 40th International Conference on Machine Learning(ICML),2023:1396-1415.
[48] Liang W,Zhao Y,She R,et al. FedSheafHN:personalized federated learning on graph-structured data[J]. arXiv preprint arXiv:2405.16056,2024.
[49] Li Y,Xu W,Wang H,et al. Personalized federated domain-incremental learning based on adaptive knowledge matching[C]//European Conference on Computer Vision(ECCV),2024:127-144.
[50] Liu Q,Sun S,Liang Y,et al. Personalized federated learning for spatio-temporal forecasting:a dual semantic alignment-based contrastive approach[C]//Proceedings of the AAAI Conference on Artificial Intelligence,Philadelphia(AAAI),2025:12192-12200.
[51] Yu H,Yang X,Gao X,et al. Personalized federated continual learning via multi-granularity prompt[C]//Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining,2024:4023-4034.
[52] Yin B,Chen Z,Tao M. Aggregation design for personalized federated multi-modal learning over wireless networks[J]. IEEE Communications Letters,2024,28(8):1850-1854.
[53] Shi J,Chen T,Zhang S,et al. Personalized quantum federated learning for privacy image classification[J]. arXiv preprint arXiv:2410.02547,2024.
[54] Dave D,Vyas K,Jayagopal J K,et al. FedGlu:a personalized federated learning-based glucose forecasting algorithm for improved performance in glycemic excursion regions[J]. arXiv preprint arXiv:2408.13926,2024.
[55] Yu P L,Kundu A,Wynter L,et al. Fed+:a unified approach to robust personalized federated learning[J]. arXiv preprint arXiv:2009.06303,2021.
[56] Sáinz Pardo Díaz J,Castrillo M,Bartok J,et al. Personalized federated learning for improving radar based precipitation nowcasting on heterogeneous areas[J]. Earth Science Informatics,2024,17(6):5561-5584.