4.1 横向联邦学习的定义

横向联邦学习也称为按样本划分的联邦学习（Sample-Partitioned Federated Learning 或 Example-Partitioned Federated Learning）[27]，可以应用于联邦学习的各个参与方的数据集有相同的特征空间和不同的样本空间的场景，类似于在表格视图中对数据进行水平划分的情况，如图 1-3 所示。事实上，“横向”一词来源于术语“横向划分（horizontal partition）”。“横向划分”广泛用于传统的以表格形式展示数据库记录内容的场景，例如表格中的记录按照行被横向划分为不同的组，且每行都包含完整的数据特征。举例来说，两个地区的城市商业银行可能在各自的地区拥有非常不同的客户群体，所以他们的客户交集非常小，他们的数据集有不同的样本 ID。然而，他们的业务模型非常相似，因此他们的数据集的特征空间是相同的。这两家银行可以联合起来进行横向联邦学习以构建更好的风控模型。确切的说，我们可以将横向联邦学习的条件总结为：

$$\mathcal{X}_i = \mathcal{X}_j,\ \ \mathcal{Y}_i = \mathcal{Y}_j,\ \ I_i \neq I_j,\ \ \forall \mathcal{D}_i, \mathcal{D}_j, i \neq j, \tag{4-1}$$

式中，$\mathcal{D}_i$ 和 $\mathcal{D}_j$ 分别表示第 i 方和第 j 方拥有的数据集。我们假设两方的数据特征空间和标签空间对，即 $(\mathcal{X}_i, \mathcal{Y}_i)$ 和 $(\mathcal{X}_j, \mathcal{Y}_j)$ 是相同的。但是我们假设两方的客户 ID 空间，即 I_i 和 I_j 是没有交集的或交集很小。

关于横向联邦学习系统的安全性的定义，我们通常假设一个横向联邦学习系统的参与方都是诚实的，需要防范的对象是一个诚实但好奇（honest-but-curious）的聚合服务器[35, 115]。即通常假设只有服务器才能使得数据参与方的隐私安全受到威胁。

文献[194]的作者提出了一种协作式的深度学习方法，其中参与方独立地训练模型并只分享参数更新的子集，这是横向联邦学习的一种特殊形式。在 2016 年，谷歌发布了一种为安卓系统手机提供模型更新的基于横向联邦学习的解决方案[12]。在谷歌提出的框架中，一部安卓手机的用户在本地更新模型参数，并将更新的模型参数上传至安卓云（Android Cloud），因此可以和其他参与方协同地训练联邦学习模型。

文献[115]的作者提出了一种在联邦学习框架下对用户模型更新或者对梯度信息进行安全聚合（secure aggregation）的方法。文献[35] 的作者提出了一种适用于模型参数聚合的加法同态加密（Additive Homomorphic Encryption，AHE）方法，能够防御联邦学习系统里的中央服务器窃取模型信息或者数据隐私。在文献[58]中，研究