项目总体框架：基于自适应联邦学习任务的可信公平区块链框架（TFchain）与面向活跃账户的区块链设计

本项目旨在构建一个 **基于区块链的自适应联邦学习系统**，结合 **面向活跃账户的区块链** 和 **可信公平区块链框架（TFchain）**，为医疗、金融等领域的多方协作提供一个高效、安全、透明且公平的联邦学习平台。该系统解决了 **算力过度消耗**、**激励机制不公平**、**数据隐私泄露** 等问题，同时引入了 **活跃账户管理** 机制，通过 **贡献度计算** 和 **智能合约** 实现公平的奖励分配和作恶节点检测。

---

1. 项目架构概述

本系统的架构由以下几个主要模块构成：

1. **区块链网络与共识机制**（面向活跃账户的区块链、PoTF共识）
2. **联邦学习模块**（基于 TFchain 支持的自适应联邦学习任务）
3. **智能合约与激励机制**（基于贡献度的奖励分配与作恶检测）
4. **账户管理与数据存储**（活跃账户表与 Verkle 树）
5. **数据安全与隐私保护**（加密、差分隐私等技术）

---

2. 系统模块与功能

2.1 区块链网络与共识机制：PoTF（Proof of Trust and Fairness）

- **区块链设计**：区块链作为整个系统的数据存储和验证层，负责记录每个节点的贡献、模型更新和交易行为，确保数据的透明性、可追溯性和不可篡改性。
- **面向活跃账户的区块链**：
  - **活跃账户表**：系统维护一个活跃账户表，记录当前处于活跃状态的节点。只有在活跃账户表中的节点，才能参与到联邦学习任务和模型更新中。通过哈希表或快速索引技术，确保活跃账户信息的快速查询和更新。
  - **账户活跃度计算**：通过基于贡献度、训练参与度等标准动态评估节点是否为活跃账户。活跃账户的状态可以由智能合约自动更新。
- **PoTF共识机制**：
  - 在每次联邦学习任务开始时，系统会选择 **贡献度最高的节点** 作为 **服务节点**，负责收集上传的模型更新并进行聚合。普通节点将进行本地训练并上传模型参数。
  - **PoTF切换机制**：根据参与方的贡献度和模型更新的质量，系统动态选择 PoW 或 PoTF 共识方式。对于贡献较高的节点，采用 PoTF 进行共识，减少算力消耗；对于贡献较低的节点，仍然可以使用传统的 PoW 机制进行验证。

2.2 联邦学习模块：TFchain支持的自适应任务

- 联邦学习任务管理

  ：

  - TFchain 提供一个 **自适应联邦学习任务** 支持框架，能够根据实际情况调整参与方的训练任务和数据共享方式。
  - **动态任务分配**：根据各节点的贡献度、资源和计算能力，系统动态调整任务分配。活跃账户将优先参与任务，而非活跃账户会根据贡献度逐步退出任务。

- 任务执行与模型更新

  ：

  - 每个参与节点进行 **本地模型训练**，并将模型更新上传至服务节点进行聚合。服务节点根据所有普通节点的贡献度计算并生成区块。
  - **模型质量评估**：服务节点在聚合过程中会评估上传模型的质量，并根据评估结果调整节点的贡献度，作为奖励和惩罚的依据。

2.3 智能合约与激励机制

- **奖励分配机制**：
  - **贡献度评估**：智能合约根据每个节点在联邦学习任务中的贡献度进行评估，包括训练时间、模型质量、参与度等。
  - **奖励机制**：根据贡献度分配奖励，贡献较高的节点获得更多奖励。奖励可以是 **代币** 或 **积分**，并可以用来作为后续任务中参与节点的资格。
  - **流动性与公平性**：服务节点的流动性保证了公平性，避免了单一节点垄断任务和奖励的现象。
- **作恶节点检测与惩罚**：
  - 通过 **投票机制** 和 **贡献度检测**，智能合约能够识别并排除恶意或不诚实的节点（如上传不合格的模型或篡改数据）。
  - 如果某个服务节点被识别为作恶节点，其将被从区块链中除名，并且其上传的模型更新不会计入全局模型中。
  - 系统也会对 **普通节点** 进行贡献度异常检测，如果节点的贡献度远低于预期（如上传的模型质量极差），该节点将被排除在奖励分配之外。

2.4 数据存储与隐私保护

- 数据存储

  ：

  - **Verkle树**：用于高效存储所有账户信息。Verkle树结合了 Merkle 树和 Trie 树的特点，能够高效地压缩数据并提高查询效率。非活跃账户的详细信息可以被存储在去中心化存储系统（如 IPFS）中，只有在必要时才通过区块链访问。

2.5 安全性与容错

- **防止攻击**：通过区块链的共识机制、智能合约的自动化管理和作恶节点检测，系统能够有效防范 **Sybil 攻击**、**分布式拒绝服务攻击（DDoS）** 和 **模型污染攻击** 等。
- **容错机制**：采用去中心化的结构，避免单点故障，确保系统的高可用性和可靠性。

---

3. 系统组成与流程

3.1 系统概述

系统由以下几个主要部分组成：

- **客户端（医疗机构）**：每个医疗机构作为客户端，本地训练模型并上传更新。每个医疗机构的模型训练基于本地的医疗数据（如 X 光片），数据不会离开本地，保护数据隐私。
- **区块链网络**：区块链记录所有节点的贡献度、模型更新以及交易信息。它确保数据和模型更新的不可篡改性和透明性，提供去中心化的信任机制。
- **聚合服务器（可选）**：通过选择历史贡献度较高的节点作为服务节点，负责收集并聚合各医疗机构上传的模型参数，生成全局模型。聚合过程中采用 **PoTF 共识机制** 来降低算力消耗并确保公正性。
- **智能合约**：智能合约自动化执行模型上传、验证、奖励分配、作恶检测等任务。通过智能合约，系统可以自动处理各节点的行为，保证训练过程的公平性，并确保奖励机制的透明性。

3.2 系统流程

1. **本地数据训练**：
   - 每个医疗机构（如医院、诊所）使用其本地的 **X 光片数据** 进行模型训练，数据始终保持本地，不会上传至中央服务器。仅有模型的更新（如梯度、权重等）被上传进行全局模型更新。
   - 在本地训练过程中，使用 **差分隐私** 和 **加密技术** 确保数据的安全性和隐私保护。
2. **模型更新上传**：
   - 各医疗机构将本地训练后的模型参数（如权重或梯度）上传至 **区块链** 网络。区块链会记录每次上传的 **哈希值** 以及时间戳等信息，以确保记录的透明性和不可篡改性。
   - 上传时，客户端通过 **智能合约** 进行验证，确保模型更新符合预定标准，并且没有恶意数据或篡改。
3. **区块链数据验证与记录**：

- 区块链使用智能合约验证每次上传的模型更新是否有效，并确保更新的来源可信。
- 每次模型更新都会生成一个包含贡献度、时间戳等信息的区块，确保上传的所有信息可追溯且不可篡改。

4. **选择聚合服务器**：
   - 每次联邦学习任务开始时，根据 **历史贡献度** 和 **信誉评分**，选择 **PoTF 共识机制** 中的服务节点。服务节点将负责收集所有上传的模型参数，进行聚合并计算各节点的贡献度。
   - 服务节点聚合模型参数并计算贡献度后，生成新区块，其中包含 **贡献度排名**、**模型参数更新** 和 **奖励分配**。
   - 贡献度的计算不仅基于模型的准确性，还考虑每个节点参与任务的频率、上传的模型质量等因素。
5. **全局模型聚合**：
   - 所有医疗机构的模型更新将通过去中心化的聚合方式进行合并。聚合方法（如加权平均、**FedAvg**）根据每个节点的贡献度来调整其在全局模型中的权重。
   - 聚合后的全局模型会返回到每个客户端，用于继续本地训练。
6. **奖励与惩罚机制**：
   - **奖励机制**：智能合约根据每个医疗机构的贡献度和模型更新质量，分配相应的奖励。高贡献的医疗机构会获得更多的奖励（如代币、积分等），鼓励更多参与。
   - **惩罚机制**：对于上传恶意数据或参与作弊的节点，系统通过 **智能合约** 自动惩罚其贡献度或将其排除在联邦学习任务之外，保证系统的公正性和可靠性。
7. **活跃账户的更新**：
   - 每次训练周期结束后，系统会更新 **活跃账户表**。活跃账户表记录了当前参与联邦学习的所有医疗机构及其贡献度。
   - 如果某医疗机构的贡献度较低，系统会将其状态更新为“非活跃”，并减少其在后续训练中的参与度。
8. **回馈与再训练**：
   - 聚合后的全局模型反馈给每个医疗机构，医疗机构继续在本地进行训练，并通过不断的更新和反馈形成一个 **闭环**，以提升全局模型的准确性和泛化能力。
   - 每轮训练结束后，医疗机构会根据上一轮的奖励和贡献度数据调整自己的训练策略，逐步提升模型的性能。