Aprendizado Federado Aplicado à Internet das Coisas

Nickolas Carlos Carvalho Silva[1] **Pedro Augusto Serafim Belo**[1]

[1]Instituto de Informática
Universidade Federal de Goiás

2024

INSTITUTO DE INFORMÁTICA UFG

UFG UNIVERSIDADE FEDERAL DE GOIÁS

Contexto

- Atualmente, existem diversos sistemas distribuídos que geram uma grande quantidade de dados diariamente e aplicam técnicas de Machine Learning (ML).
- Em geral, existem pequenos volumes de dados distribuídos em uma grande quantidade de entidades, sejam corporações ou indivíduos.
- Além disso, o modo como os dados são utilizados no treinamento de ML pode infringir algumas leis voltadas para a privacidade dos usuários (ex.: LGPD, GDPR, CCPA).

Contexto

- Os métodos de Machine Learning terão que se adequar a um ambiente em que os dados existem, porém estão espalhados em um grande sistema distribuído.
- A transferência de dados de um sistema distribuído para uma entidade central que irá treinar os modelos pode não ser viável, devido a:
  - Questões relativas à privacidade e regulações.
  - Alto custo da transferência de grandes quantidades de dados em determinadas redes.
- Diante disso, o **Aprendizado Federado** (*Federated Learning*) surge como uma solução para o problema de treinar modelos de ML em ambientes que apresentam alta fragmentação de dados sem violar as regulações voltadas à privacidade dos usuários.

Visão Geral

O que é?

O **Aprendizado Federado** é um método em que múltiplos dispositivos colaboram para o treinamento de um **modelo global** sem a necessidade de compartilhar seus **dados de treinamento locais**.

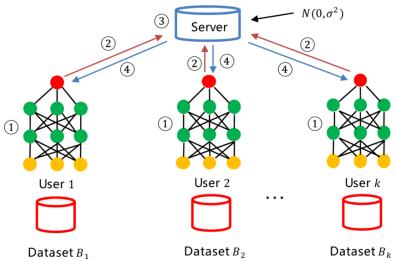


Figure 1: Fonte: *Huang et al.*

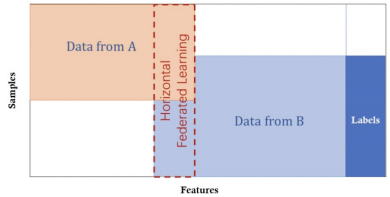
Vantagens

- De acordo com [Lim et al., 2020], em comparação com as abordagens convencionais, o Aprendizado Federado apresenta as seguintes vantagens:
  - → **Uso altamente eficiente da largura de banda da rede**
  - → **Privacidade**
  - → **Baixa latência**

Aprendizado Federado Horizontal

O que é?

O **Aprendizado Federado Horizontal** é uma configuração de AF que pode ser aplicada em cenários nos quais conjuntos de dados em locais diferentes compartilham características/*features* semelhantes, mas diferentes instâncias/*samples*.



Aprendizado Federado Horizontal

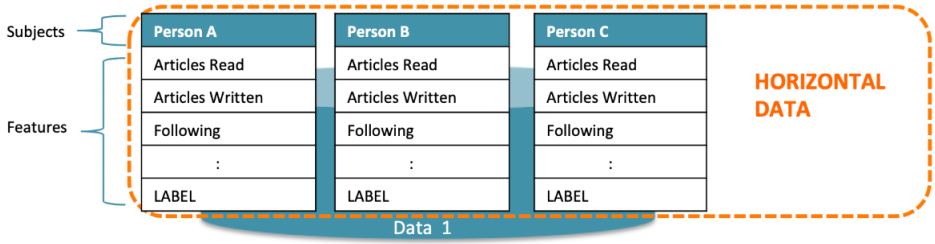


Figure 2: Ilustração do Aprendizado Federado Horizontal

Algoritmos de Média Federada

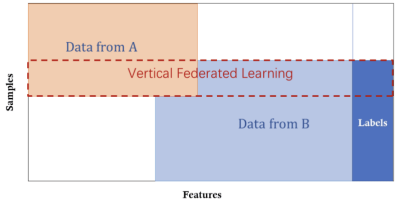
Algoritmos de Média Federada têm a finalidade de agregar os modelos locais dos participantes convergindo-os para um modelo global. O algoritmo de média federada FedAvg, empregado por [McMahan et al. 2016a], pode ser descrito da seguinte forma:

1. O servidor inicia um modelo global;
2. Em cada rodada, é selecionada uma fração ρ dos participantes;
3. O modelo global é enviado para os participantes selecionados;
4. Os participantes realizam o treinamento do modelo com os dados locais;
5. Cada participante envia o novo modelo local para o servidor agregador;
6. O servidor ajusta o modelo global a partir da média dos parâmetros dos novos modelos recebidos.

Aprendizado Federado Vertical

O que é?

O **Aprendizado Federado Vertical** é uma configuração de AF que pode ser aplicada em cenários nos quais conjuntos de dados em locais diferentes compartilham o mesmo espaço de amostras/*samples*, mas características/*features* distintas.



Aprendizado Federado Vertical

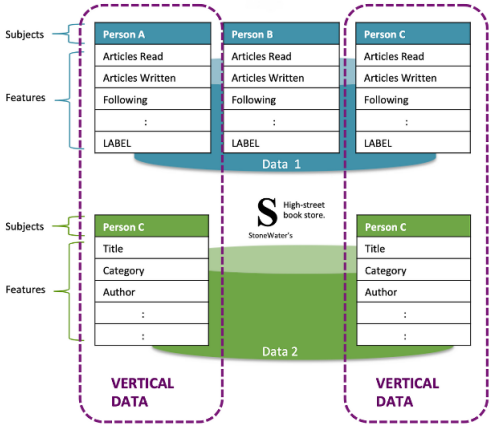


Figure 3: Ilustração do Aprendizado Federado Vertical

Referências

- Aprendizado Federado aplicado à Internet das Coisas – XXXIX Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos (SBRC 2021)
- Understanding the types of federated learning – OpenMined
- What is federated averaging (FedAvg)? – Educative.io
- DP-FL: a novel differentially private federated learning framework for the unbalanced data