

<p style="text-align: center;">Universidade de Pernambuco Programa de Pós-Graduação em Engenharia da Computação (PPGEC)</p>
<p style="text-align: center;">Proposta de Dissertação de Mestrado</p>
<p>Área: Inteligência Computacional</p>
<p>Título: FedP2P++: Estratégias Avançadas de Colaboração Peer-to-Peer em Aprendizado Federado</p>
<p>Orientador – Bruno José Torres Fernandes (bjtf@ecomp.poli.br)</p>
<p>Descrição – O avanço da tecnologia e a coleta massiva de dados têm impulsionado o desenvolvimento de algoritmos de aprendizado de máquina cada vez mais complexos e eficientes [1]. No entanto, muitas vezes esses dados são distribuídos em diferentes locais, o que torna difícil a tarefa de treinar um modelo com essas informações [2]. Além disso, o compartilhamento desses dados pode violar a privacidade das pessoas ou empresas envolvidas [3].</p> <p>O aprendizado federado é um paradigma recente [4] que permite o treinamento de modelos de aprendizado de máquina em conjuntos de dados distribuídos e descentralizados, sem a necessidade de compartilhar os dados diretamente entre os participantes. O aprendizado federado permite então que múltiplas entidades (por exemplo, dispositivos móveis, sensores, servidores em diferentes locais) colaborem para treinar um modelo sem compartilhar seus dados brutos. Recentemente, tem ganhado destaque uma nova abordagem chamada aprendizado federado peer-to-peer, que incorpora mecanismos de colaboração direta entre clientes. Contudo, mesmo com resultados animadores, diversas questões permanecem em aberto: qual o impacto do refinamento entre pares em diferentes domínios de aplicação? Como garantir eficiência em cenários com milhares de clientes e redes heterogêneas? E, principalmente, como estender essa abordagem para além de um simples wrapper, investigando suas propriedades teóricas, variantes adaptativas e integração com outras estratégias avançadas?</p> <p>Este projeto de mestrado propõe investigar essas lacunas por meio da exploração sistemática de estratégias peer-to-peer em aprendizado federado. O objetivo é desenvolver, analisar e avaliar arquiteturas baseadas em peer-to-peer que superem os desafios impostos pela heterogeneidade extrema dos dados, viabilizando modelos mais robustos, rápidos e equitativos. Diferente das abordagens puramente centralizadas, nossa proposta busca compreender e explorar os efeitos da colaboração distribuída com ênfase em convergência, bias, fairness e escalabilidade, contribuindo tanto com fundamentos teóricos quanto com aplicações práticas em ambientes reais.</p> <p>Importante destacar que o referido projeto está alinhado a bolsa CAPES-Humboldt do professor proponente, que é desenvolvida em parceria com a Leibniz Universität Hannover, Alemanha.</p>

Referências Bibliográficas

1. NGUYEN, G. T. et al. Machine learning and deep learning frameworks and libraries for large-scale data mining: a survey. *Artificial Intelligence Review*, v. 52, p. 77–124, 2019.
2. TULADHAR, A. et al. Building machine learning models without sharing patient data: A simulation-based analysis of distributed learning by ensembling. *Journal of biomedical informatics*, p. 103424, 2020.
3. ZHANG, T.; HE, Z.; LEE, R. B. Privacy-preserving machine learning through data obfuscation. *ArXiv*, abs/1807.01860, 2018.
4. MCMAHAN, H. B. et al. Communication-efficient learning of deep networks from decentralized data. In: *International Conference on Artificial Intelligence and Statistics*. [S.l.: s.n.], 2016.