Survey em IA e Segurança

Davi Iury, Esther Martins, Lucas Pinheiro, Rafael Porto, and Théo Araújo

Universidade Federal do Ceará

Abstract. Nós resume as coisas aqui.

Keywords: IA · Segurança · Machine Learning.

1 Introdução

IA é muito popular. Porém, precisamos de muitos dados para treinar modelos. Como podemos arranjar esses dados? Mais especificamente, como podemos arrumar esses dados de forma que não infrijamos leis de privacidade de dados? Como privacidade de dados vem se tornando um conceito cada vez mais em voga, Novas formas de treinar modelos e obter dados vem surgindo. Nesta survey, falaremos de:
Federated learning (analisar os dados locamente e mandar os resultados de volta de forma criptografada)
Differential privacy (é uma técnica que visa proteger a privacidade dos usuários por meio da adição de ruído nos dados sendo analisados)
Machine Unlearning (esquecer dados de usuários que foram usados para treinar modelos de forma que isso não prejudique o aprendizado do algoritmo).

1.1 Contextualização

Deixei aqui para ser o template inicial. Quando forem escrever, É legal dar enter a cada oração Para que fique dividido direito e fique fácil de ler.

2 Caracterização Ferramental

2.1 Machine unlearning

woow

2.2 Differential privacy

woow

2.3 Federated learning

Contexto. É comum que algoritmos clássicos de aprendizado de máquina mantenham centralizados os dados a serem usados para treinamento. Isso se deve ao fato de que, frequentemente, os dados encontram-se dispersos em ¨ilhas de dados¨[1], então, é necessário que seja feito um trabalho de captação e agrupamento em um servidor para que o uso em treinamento seja possível. No entanto, caso não seja feita de forma adequada, essa centralização facilita o vazamento de dados sensíveis. Em vista dessa situação, diversas regulações vem sendo impostas com relação à captação e ao uso de dados para treinamento de modelos, tornando o uso de técnicas de aprendizado de máquina centralizado de difícil implementação prática. Nesse contexto, a aplicação do federated learning possibilita com que o treinamento seja feito de forma local, não-centralizada, de forma que os dados de um usuário específico mantém-se somente no seu dispositivo local.
Definindo. Em sua essência, o federated learning é uma técnica de aprendizado distribuído, ou seja, ao invés de consolidarmos dados de usuário em um servidor central para treinar um modelo, haverão focos locais de treinamento. Assim, evitando a captação de dados sensíveis, o servidor envia um modelo de treinamento para cada dispositivo individual, mantendo a fase de treinamento como uma etapa local. Cada dispositivo somente retornará ao servidor central seu modelo treinado localmente e atualizará o seu modelo interno conforme as atualizações no modelo global. Federated learning, dessa forma, garante que dados locais não possam ser vazados e, a fim de não violar leis gerais de proteção de dados, a troca de parâmetros entre clientes locais e o servidor para a geração de um modelo global é feita através de mecanismos criptografados.[2]
Tipos.
1. Com relação à partição dos dados.
a. Sistemas que fornecem os dados tem usuários muitos diferentes, mas features parecidas b. Sistemas tem features muito diversas, mas usuários muito parecidos c. Sistemas fornecem dados muito incompatíveis (features e usuários diferentes) ou insuficientes.
2. Com relação à mecanismos de privacidade

3. Com relação ao modelo de ML aplicado

4. Com relação ao método de solucionar heterogênidade.

Aplicações práticas.

3 Desafios

[1] to-do referencia.
[2] botar uma figura aqui