



Inteligência Artificial

- Algoritmos estão tomando cada vez mais decisões
- Estas decisões são feitas automaticamente
- Até que ponto podemos confiar neste modelos?

Aplicações



Diagnóstico Médico



Aprovação de Crédito



Veículos Autônomos



Decisões Judiciais



Contratação de Funcionários

1. Questões Legais

Legislação obriga



2.Reputação

AI toma uma decisão preconceituosa



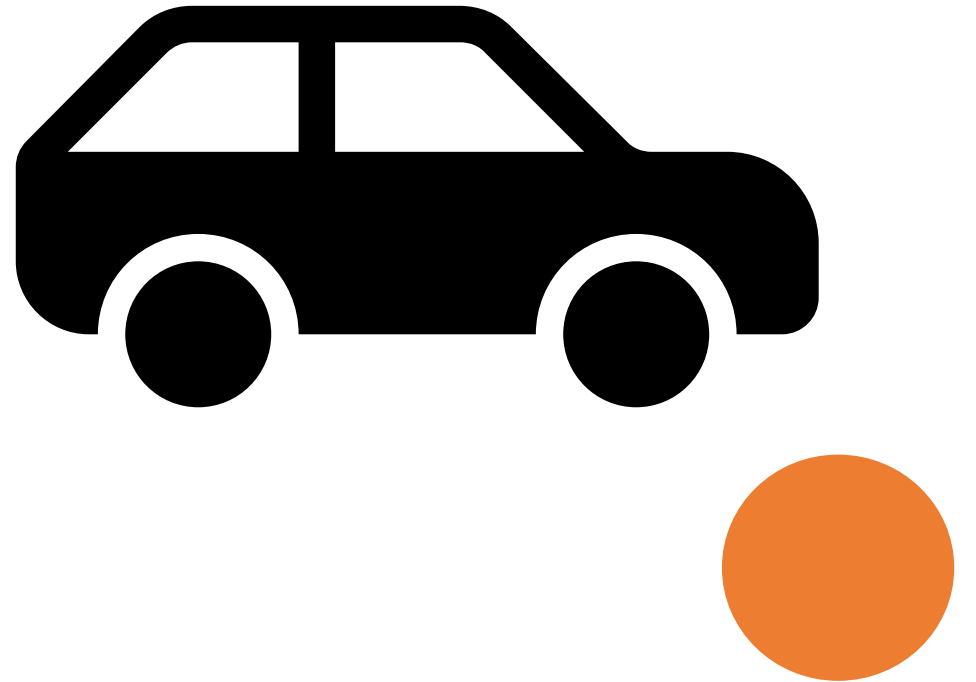


Cinco vezes em
que IA foi
discriminatória

- <https://www.newscientist.com/article/2166207-discriminating-algorithms-5-times-ai-showed-prejudice/>

3. Segurança Jurídica/Financeira

- Um veículo autônomo provoca um acidente
- Como planejar uma defesa?





Self-Driving Uber Car Kills Pedestrian in Arizona

<https://www.nytimes.com/2018/03/19/technology/uber-driverless-fatality.html>

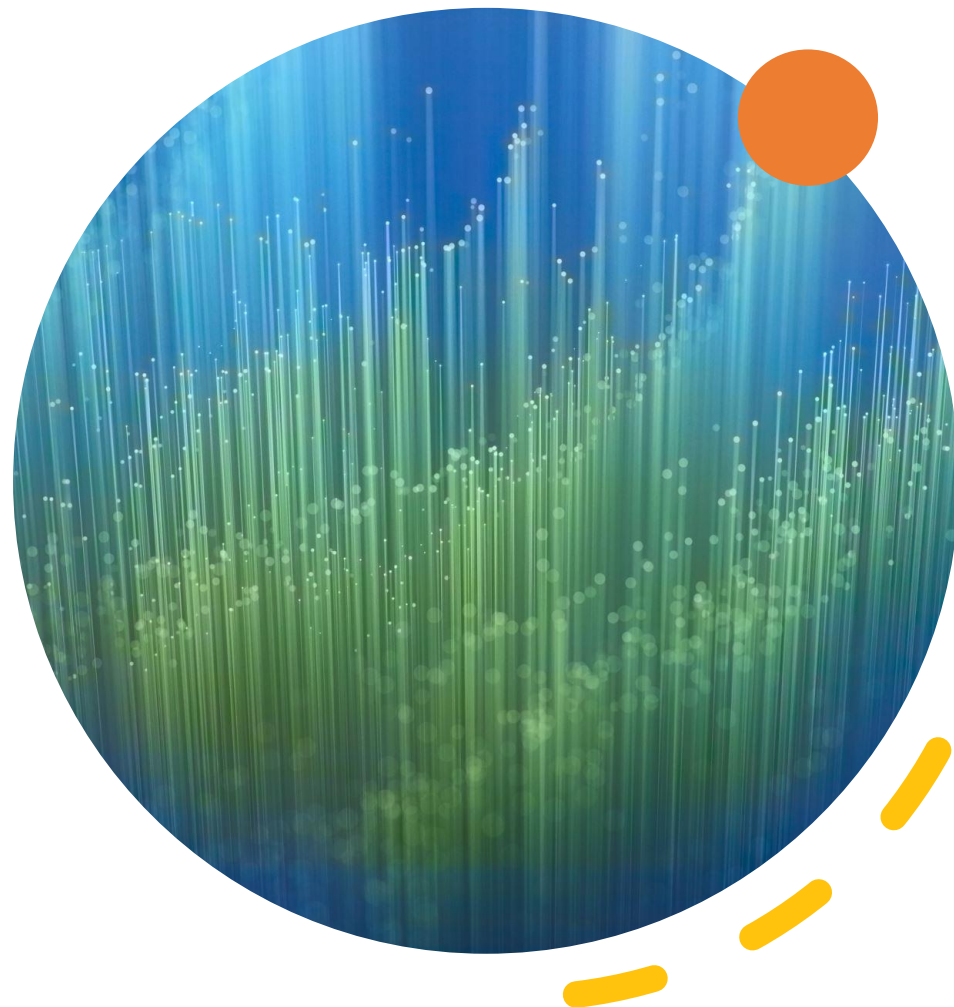
4. Depuração



NORMALMENTE O MODELO É TESTADO
EM TERMOS DE PRECISÃO E OUTRAS
MÉTRICAS. MAS ISSO É SUFICIENTE?



ENVIESAMENTO NO TREINO



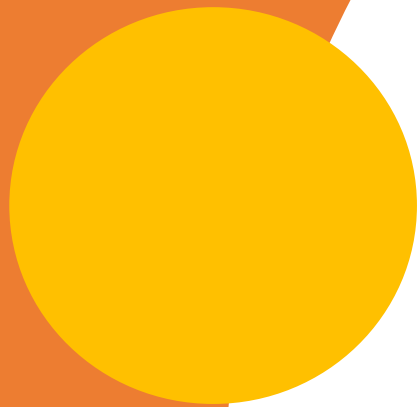
5.Prevenção de Ataques

Um modelo pode ser "hackeado"



One Pixel Attack for Fooling Deep Neural Networks

<https://arxiv.org/pdf/1710.08864.pdf>



6. Questão "pessoal"

Executivo quer entender