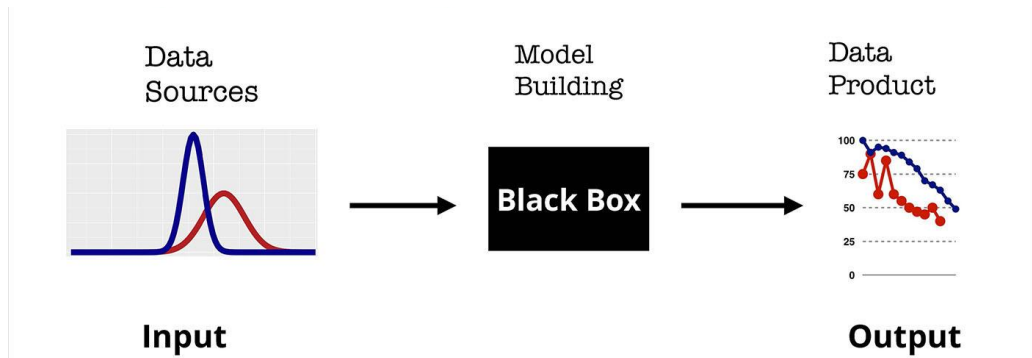


Máquinas Preconceituosas: Machine Learning e Ciências Sociais

Aluno: Matheus Feliciano (mbdf)

Algoritmos de Machine Learning sendo usados em alta escala para análise da sociedade

- Nossa sociedade é preconceituosa o que pode trazer um viés discriminatória para os dados
- Os algoritmos geralmente são caixas pretas que tem como entrada dados de fontes diversas e dão como resultado soluções ou produtos baseados na interpretação dos dados pela máquina





Como os resultados podem ter um viés preconceituoso?

- Rótulos dados por seres humanos aos dados
- Dados de treinamento não suficientes ou tendenciosos
- Impacto na representatividade de certos grupo
- Muitas vezes os resultados discriminatórios são dados por associação de padrões
- Sistema cíclico se houver a construção de uma nova iteração com resultados tendenciosos



Possíveis danos causados por esses sistemas

- Dano Alocativo: quando um sistema aloca ou não aloca certo grupo para um oportunidade ou recurso
 - Estereotipação, Reconhecimento, Difamação, Sub-representação e Exoneração
- Dano Representativo: quando um sistema reforça as subordinações de certos grupos fazendo uma alocação má distribuída de acordo com raça, gênero, religião, etc

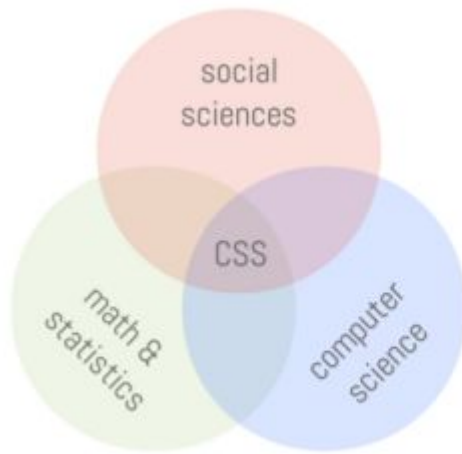
Monitoramento de modelos

- Modelos: Predição, Explicação e Exploratório
- Predição: tentar prever comportamento através dos dados
- Explicação: tentar entender o que causou os resultados que os dados descrevem
- Exploração: tentar entender novos padrões sociais a partir dos dados coletados
- Monitorar modelos para que não haja associações com cunho discriminatório de forma a verificar como os dados estão sendo interpretados pelo modelo



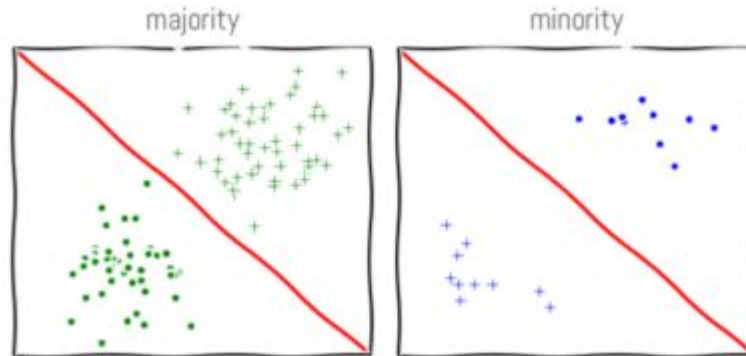
O papel do cientista social

- Acompanhamento do processo da geração dos modelos e dos resultados
- Conscientização da comunidade que a interdisciplinaridade é essencial em temas como este e para conseguir evitar esses danos e problemas
- Para modelos exploratórios ter uma atenção maior na geração de novas formas de avaliação



Possíveis soluções

- Uma análise mais específica das minorias em um conjunto de dados
- Separar em pequenos clusters de minorias e avaliar cada um deles de forma a interpretar os dados dessas minorias de forma diferente das majorias





Conclusão

- Primeiramente devemos ter concepção que haverá implicitamente um preconceito, já que ele existe na nossa sociedade, e que devemos desafiar os estereótipos de minorias, mesmo que dê mais trabalho
- Em segundo lugar, se quisermos usar outros métodos para responsavelmente e justamente analisar esses dados precisamos de mais envolvimento interdisciplinar da comunidade como um todo para o melhor entendimento da ciência