

O Enviesamento Intrínseco e Escondido em Big Data



Um resumo dos artigos *Big Data's Disparate Impact*, de Solon Barocas & Andrew D. Selbst e *The Hidden Biases in Big Data*, de Kate Crawford

Henrique Caúla | lhnc
Luís Henrique Nascimento | lhns

Big Data em Interações Humanas



Estudos sociais



Recrutamento em
empresas

Por natureza, haverá distinção

Garbage in, garbage out



Um algoritmo é apenas tão bom quanto os dados que trabalha

Solon Barocas, Andrew D. Selbst

Discriminação

Etapas vulneráveis a viés humano

Classificação

Labeling

Spam

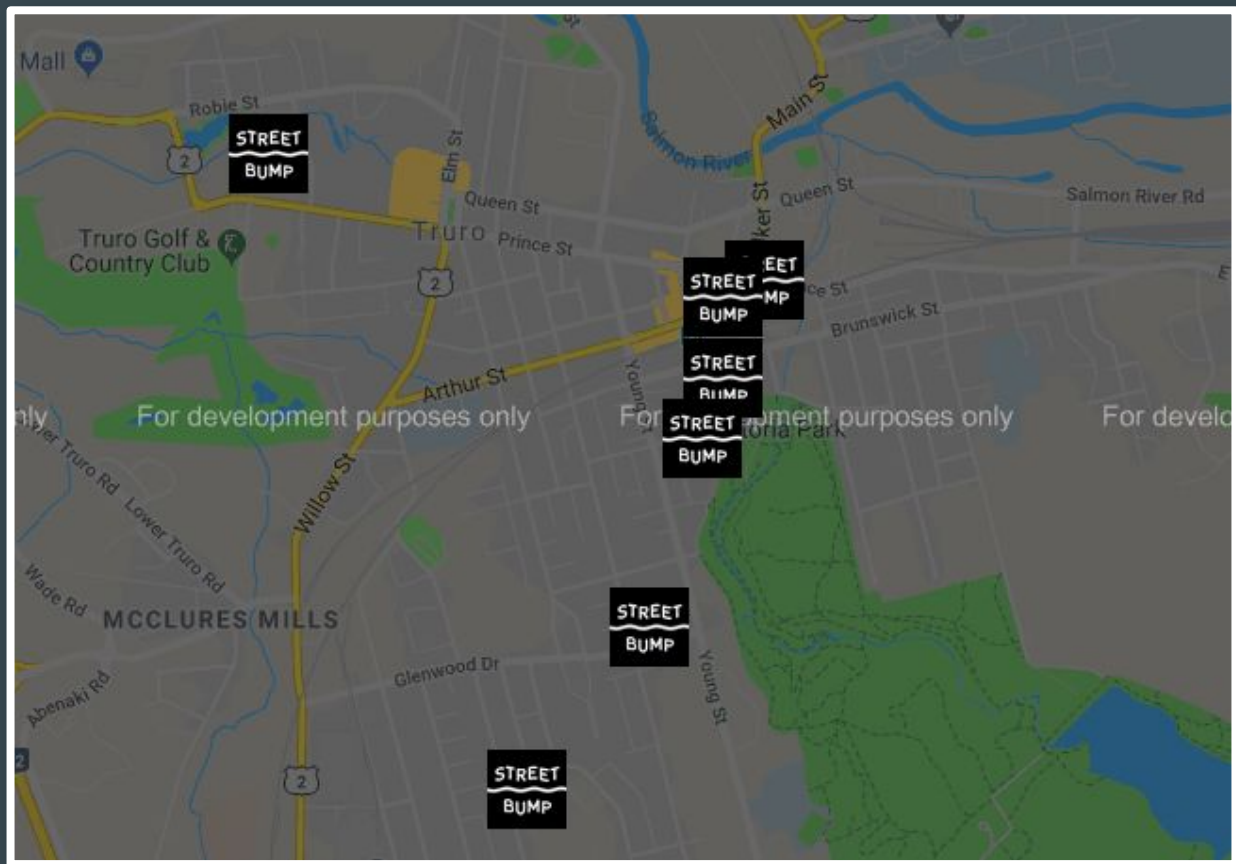
distinguível
binário
não controverso

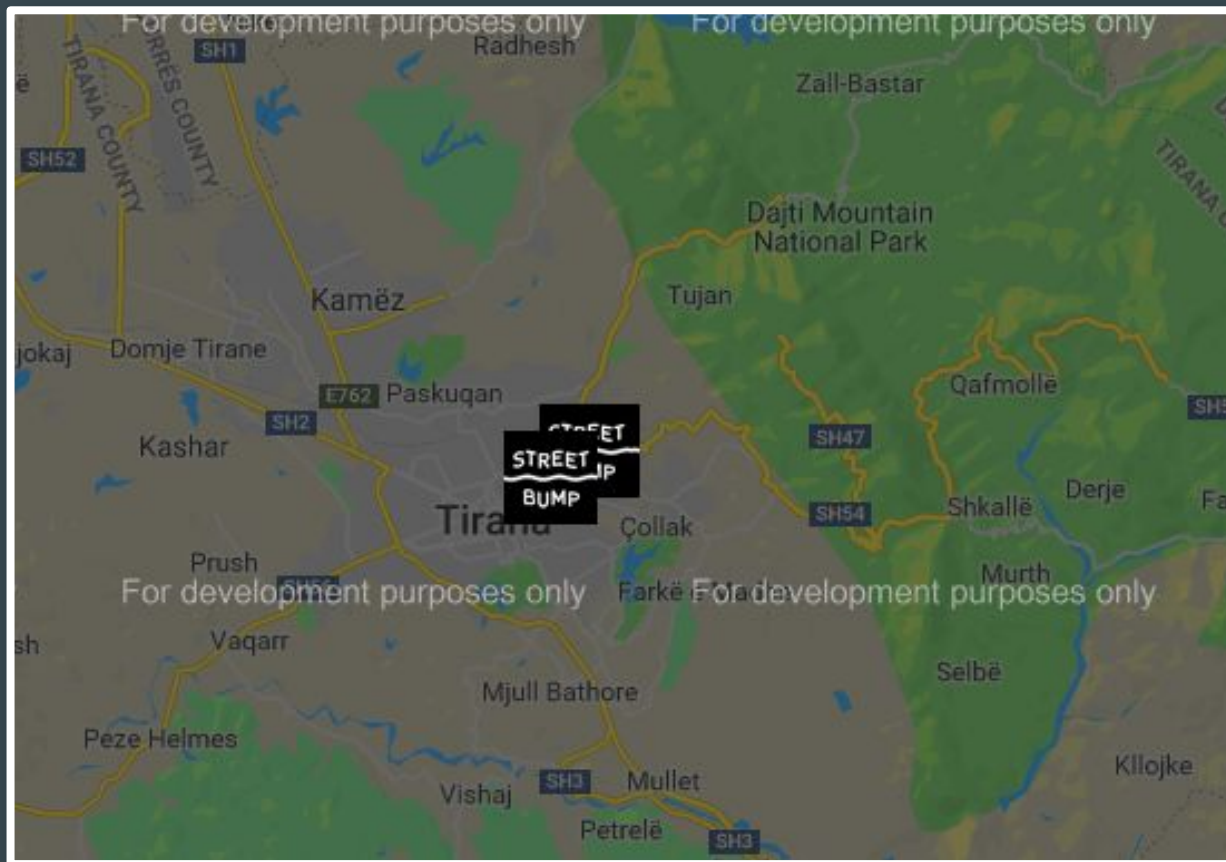
Recrutamento

não determinístico
mais de uma opção
sujeito a discriminação

Sub/Sobre Representação

Mais ou menos classes





Seleção de atributos

redlining

Local de residência

Mais perto -> mais engajado
Naturalmente há segregação

Curso superior

Curso melhor -> melhor performance
Naturalmente há segregação

Masking

Uso explícito ou implícito de
discriminação



Antecedente criminal



Teste de personalidade
cultural

Suscetibilidade

Intencional

Conhecimento do viés
Reconhecimento da possibilidade de viés
Resultado replicável

Alternativas

Pode-se utilizar outro meio?
Pode-se modificar os dados?

Próximos passos

Estudo técnico de times legais

Vulnerabilidade em cada passo

Detecção de vulnerabilidade

Estudo dos analistas para reduzir a
possibilidade de discriminação

Equilíbrio - precisão x discriminação

Reduzir ao mínimo a qualidade dos
algoritmos

Obrigado!

Um resumo dos artigos *Big Data's Disparate Impact*, de Solon Barocas & Andrew D. Selbst e *The Hidden Biases in Big Data*, de Kate Crawford

Henrique Caúla | lhtc
Luís Henrique Nascimento | lhns