TECS
Comput{ação social}

Inteligências artificiais, preconceitos reais

Inteligência artificial

TECS
Comput{ação social}

"A teoria e o desenvolvimento de **sistemas computacionais capazes de executar tarefas que normalmente requerem inteligência humana**, como percepção visual, reconhecimento de fala, tomada de decisão e tradução entre idiomas" **- Google Dictionary**

"I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be **able to speak of machines thinking without expecting to be contradicted**" **- Alan Turing, 1947**

Inteligência artificial

TECS
Comput{ação social}

- No fundo, algoritmos [de *machine learning*] são simplesmente **ferramentas** que **“aprendem”** (isto é, aumentam sua precisão) **com um grande volume de dados e fornecem algum tipo de resposta otimizada** - como um ranking ou uma avaliação - de acordo com o **procedimento pré-programado**.

TECS
Comput{ação social}

A importância dos dados

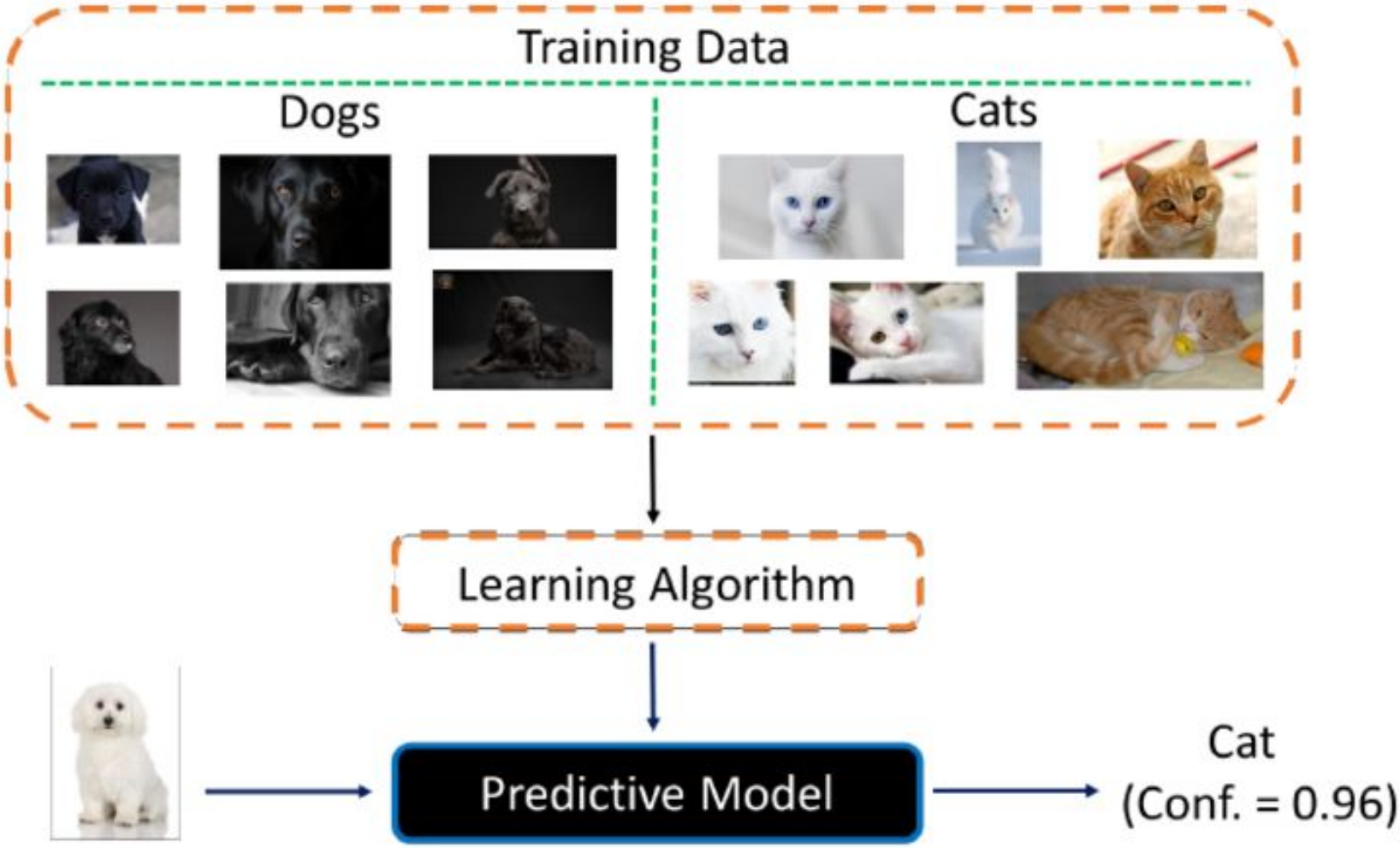
- Uma vez que a grande maioria das tecnologias conhecidas como **“inteligências artificiais” são feitas em cima de análise exploratória de dados**, estes determinam em muito o comportamento dessas ferramentas

- Há uma **diferença** fundamental no tratamento de dados sobre **ciências naturais** e os sobre **ciências humano-sociais**

Desafios éticos

- Relacionados aos **dados utilizados como inputs** para os algoritmos

- Relacionados ao **funcionamento interno e projeto do algoritmo**


Training Data
Dogs
Cats
Learning Algorithm
Predictive Model
Cat
(Conf. = 0.96)

Possíveis problemas

TECS
Comput{ação social}

- **Dados mal selecionados:** escolhas humanas sobre **o uso de certos dados (e não outros)** podem resultar em escolhas discriminatórias.
- **Dados incompletos, incorretos ou desatualizados:** falta de acurácia ou existência de lacunas nos dados; falta de rigor técnico na coleta.
- **Viés de seleção:** dados coletados não são **representativos da população**.
- **Perpetuação e promoção não intencionais de viéses históricos:** dados refletem **resultados passados**, que podem ser discriminatórios.

NEXO

ASSINE CONTA

EXPRESSO

O desafio de ensinar máquinas a entender o mundo sem reproduzir preconceitos

Murilo Roncolato 22 Ago 2017 (atualizado 22/Ago 16h47)

Pesquisadores identificaram que sistemas 'inteligentes' passaram a ligar a ação de 'cozinhar' muito mais a mulheres que a homens

Cornell University Library

We gratefully acknowledge support fro the Simons Foundatio and member institution

arXiv.org > cs > arXiv:1707.09457

Search or Article ID All fields

(Help | Advanced search)

Computer Science > Artificial Intelligence

Men Also Like Shopping: Reducing Gender Bias Amplification using Corpus-level Constraints

Jieyu Zhao, Tianlu Wang, Mark Yatskar, Vicente Ordonez, Kai-Wei Chang

(Submitted on 29 Jul 2017)

Download:

- PDF
- Other formats

(license)

Current browse context: cs.AI

TECS Comput{ação social}

Possíveis problemas

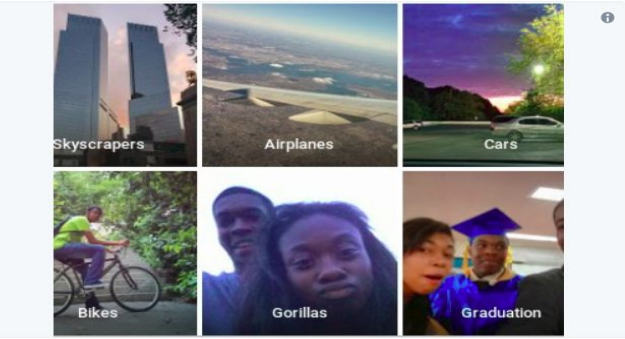
TECS
Comput{ação social}

- **Desconsideração dos dados de input:** algoritmos que não levam em conta **vieses históricos ou pontos cegos dos dados utilizados para treinamento**.
- **Diminuição do fluxo de informações:** o uso de informações detalhadas sobre o usuário é usado para inferir suas preferências, **restringindo o fluxo de informações que ele recebe**, causando uma desigualdade de oportunidades e inclusão.
- **Correlação e causação** são coisas **diferentes!**

EL PAÍS TECNOLOGIA

Google conserta seu algoritmo “racista” apagando os gorilas

Google Photos confundia pessoas negras com macacos. Este patch mostra a opacidade dos algoritmos

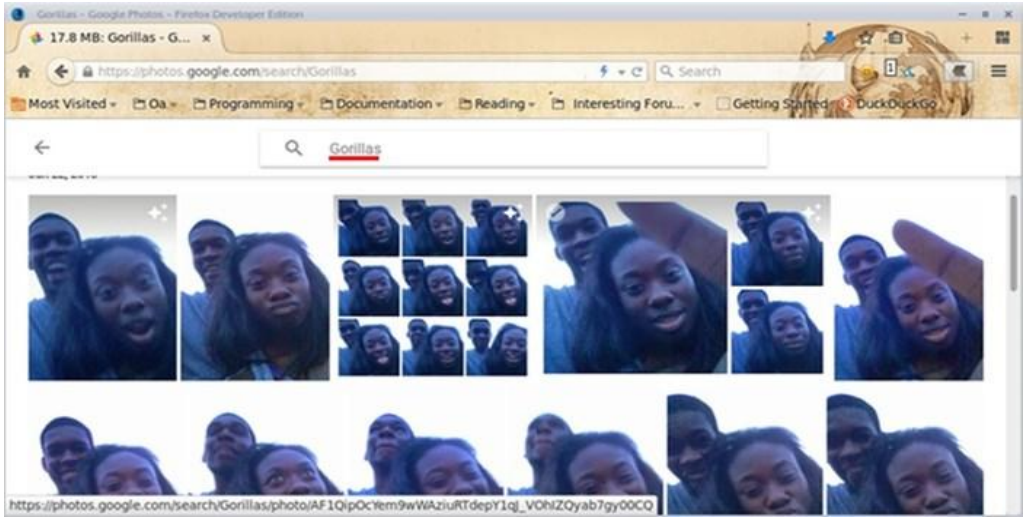


jackyalciné is about 40% into the IndieWeb.
@jackyalcine

Google Photos, y'all fucked up. My friend's not a gorilla.

22:22 - 28 de jun de 2015

2.280 3.591 pessoas estão falando sobre isso



Desenho de algoritmos

TECS
Comput{ação social}

Microsoft made a chatbot that tweets like a teen

By Jacob Kastrenakes | @jake_k | Mar 23, 2016, 10:26am EDT

SHARE

MOST REAI

Microsoft is trying to create AI that can pass for a teen. Its research team launched a chatbot

"O chatbot foi criado numa colaboração entre a **equipe de Tecnologia e Pesquisa da Microsoft e sua equipe do Bing**. Eles dizem que as habilidades de conversação de Tay foram **construídas por "mineração de dados públicos relevantes"** e combinadas com a contribuição da equipe editorial, "incluindo comediantes de improvisação". **Supõe-se que o bot aprenda e melhore à medida que fala com as pessoas**, então, teoricamente, ele se tornará mais natural e melhor entendendo a entrada ao longo do tempo."

The Verge, 23/03/2016

Desenho de algoritmos

TECS
Comput{ação social}

VERGE TWEET SHARE

The Verge, 24/03/2016

Mathwashing

TECS
Comput{ação social}

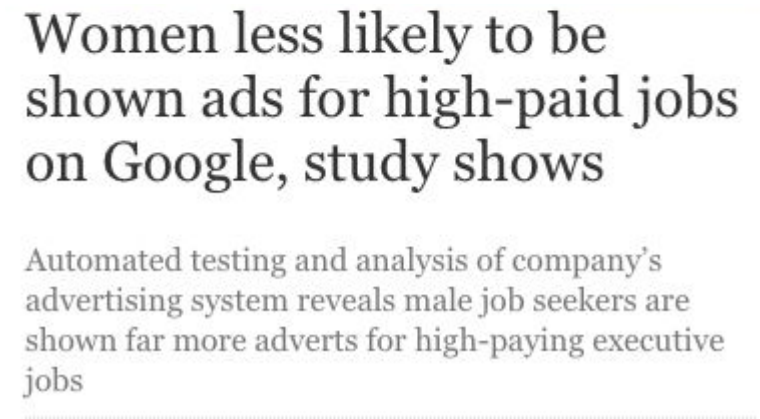
- *Mathwashing* pode ser pensado como o **uso de termos matemáticos (algoritmo, modelo, etc.) para encobrir uma realidade mais subjetiva**.
- Algoritmos complexos são **programados por humanos que**, para todos os fins e propósitos, **são imperfeitos**.
- Seu funcionamento é baseado na **análise de dados do passado**, o que faz com que tendam, por sua própria natureza, a **repetir erros humanos e perpetuá-los** através de *loops* de feedback.

Mathwashing

TECS
Comput{ação social}

Acidental

Boas intenções + falta de conhecimento + expectativas inocentes



Women less likely to be shown ads for high-paid jobs on Google, study shows

Automated testing and analysis of company's advertising system reveals male job seekers are shown far more adverts for high-paying executive jobs

The Guardian

Proposital

Como as pessoas não questionam decisões de um algoritmo "neutro", abusa-se dessa crença.

Former Facebook Workers: We Routinely Suppressed Conservative News

Michael Nunez
5/09/16 9:10am
Filed to: FACEBOOK
91

Gizmodo

Mathwashing

TECS
Comput{ação social}

EL PAÍS ATUALIDADE

INTELIGÊNCIA ARTIFICIAL ›

O inconsistente e perigoso 'radar gay'

Estudo polêmico cria algoritmo que supostamente identifica gays através de inteligência artificial

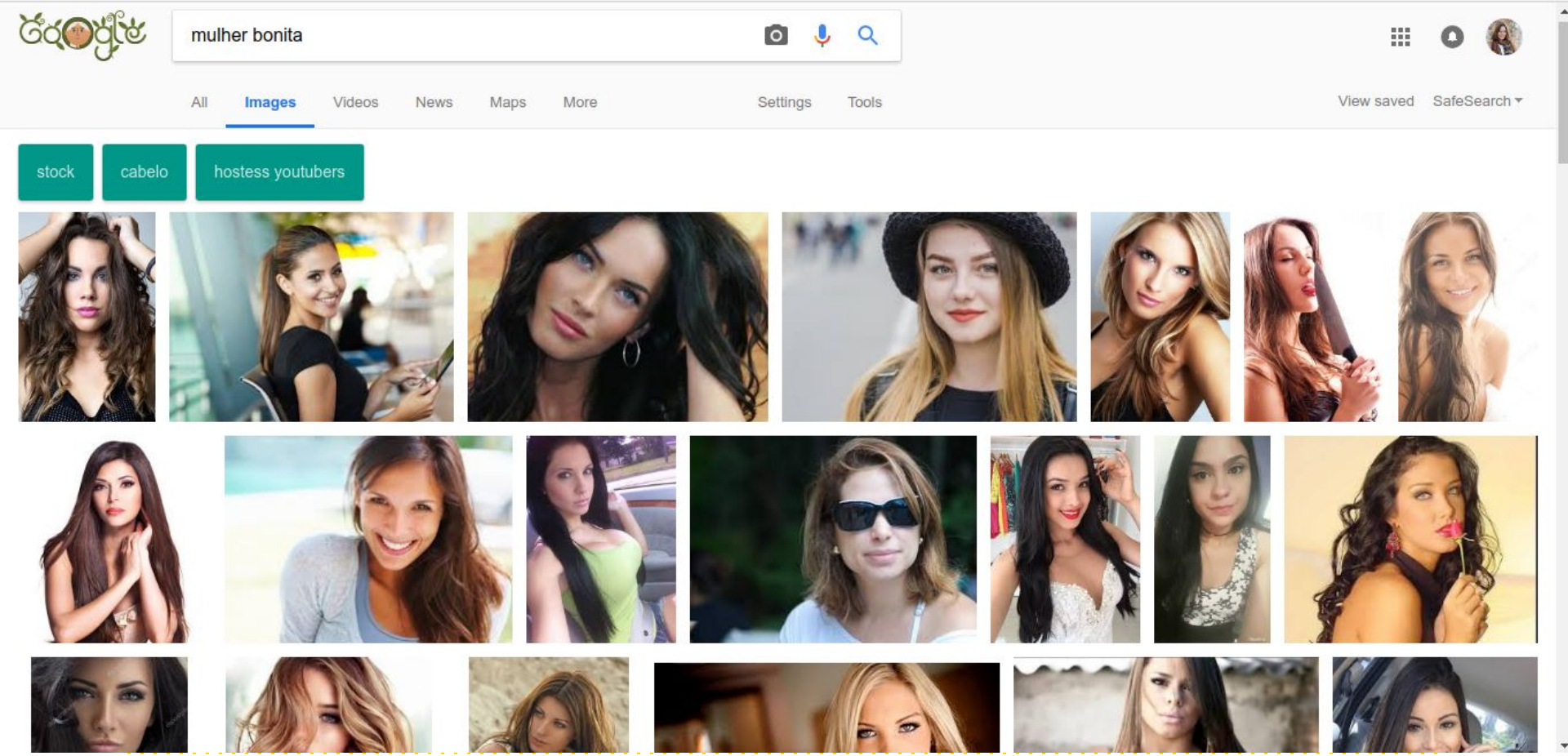
"A inteligência artificial pode determinar se alguém é heterossexual e homossexual através de seus traços faciais. Essa é a polêmica conclusão de um estudo realizado por pesquisadores da Universidade de Stanford, segundo o qual **uma IA pode distinguir se um homem é gay em até 91% dos casos e se uma mulher é lésbica em 83%**, uma porcentagem sensivelmente superior ao olho humano, que acerta 61% e 54% das vezes respectivamente, de acordo com a pesquisa."

El País, 12/09/2017

Ciclo vicioso

- **Os algoritmos de *machine learning* não apenas analisam nossas crenças e comportamento, mas também os moldam.**

Estudo de Robert Epstein: pessoas utilizavam um simulador de buscadores, que fornecia as páginas em ordens diferentes para cada grupo. Após serem expostas aos resultados dos buscadores, **as crenças das pessoas eram fortemente afetadas pelo conteúdo das páginas no topo das buscas**.


mulher bonita
All
Images
Videos
News
Maps
More
Settings
Tools
View saved
SafeSearch
stock
cabelo
hostess youtubers


French
Turkish
Portuguese
Detect language
Portuguese
English
Spanish
Translate
O bir dóktor
12/5000
He is a doctor
Suggest an edit
French
Turkish
Portuguese
Detect language
Portuguese
English
Spanish
Translate
O bir hemsire
13/5000
She is a nurse
Suggest an edit

TECS
Comput{ação social}

E agora?

E agora?

TECS
Comput{ação social}

Em um relatório do GT sobre Big Data da Casa Branca, foram delineados os seguintes pontos:

- Apoiar pesquisas que busquem eliminar a discriminação algorítmica, **construindo sistemas justos e responsáveis** [*accountable*].
- Estimular os participantes do mercado a desenvolverem **sistemas algorítmicos transparentes e com mecanismos de responsabilidade**, como a habilidade dos sujeitos de **corrigir dados incorretos sobre si** e de **entrar com recursos contra decisões algorítmicas**.

E agora?

- Promover pesquisas na academia e na indústria sobre **auditoria e testes externos de algoritmos** para garantir que as pessoas estejam sendo tratadas de maneira justa.
- Aumentar a **participação do público** na ciência da computação e de dados, incluindo oportunidades para aumentar a **fluência digital do cidadão médio**.
- Considerar o papel do governo e do setor privado em determinar as **regras de como os dados podem ser utilizados**.

Para saber mais...

- https://www.mathwashing.com/
- ESTADOS UNIDOS. **Big Data:** A report on Algorithmic Systems, Opportunity, and Civil Rights, Washington, mai. 2016.
- NOBLE, Safyia Umoja. **Algorithms of Opression:** How Search Engines Reinforce Racism. Nova Iorque: NYU Press, 2018.
- O'NEILL, Cathy. **Weapons of Math Destruction:** How Big Data Increases Inequality and Threatens Democracy. Nova Iorque: Crown Publishers, 2016.

TECS
Comput{ação social}

Participe!

Vem pro Tecs! Você pode nos encontrar em:

t.me/tecsusp

fb.com/tecs.usp

www.ime.usp.br/~tecs/