Prospectando o futuro

Inteligência Artificial e Machine Learning como apoio de notários e registradores

ONR e Centrais notariais e de Protesto

Marco Almada

lawgorithm*}

Sobre o palestrante

Marco Almada. Prospectando o futuro.

- Pesquisador no Instituto Universitário Europeu
  - Regulação da IA
- Formação
  - Mestre em Engenharia da Computação (Unicamp)
  - Bacharel em Ciência da Computação (Unicamp)
  - Bacharel em Direito (USP)

Agenda para discussão

1. Do que falamos ao falar de inteligência artificial?
2. Oportunidades e riscos da inteligência artificial
3. Relevância para as atividades notariais e registrais
4. Inteligência artificial e o Provimento 88/2019

A inteligência artificial hoje

Aplicações e o estado da arte

The Genealogy of Ideology: Predicting Agreement and Persuasive Memes in the U.S. Courts of Appeals

Shivam Verma
Department of Mathematics
Courant Institute of Mathematical Sciences, New York University
New York, New York 10003
sv1239@nyu.edu

Adithya Parthasarathy
Department of Computer Science
Courant Institute of Mathematical Sciences, New York University
New York, New York 10003
ap4608@nyu.edu

Daniel L. Chen
Institute for Advanced Study, Toulouse School of Economics
21 alle de Brienne
Toulouse 31015, France
daniel.li.chen@gmail.com

ABSTRACT

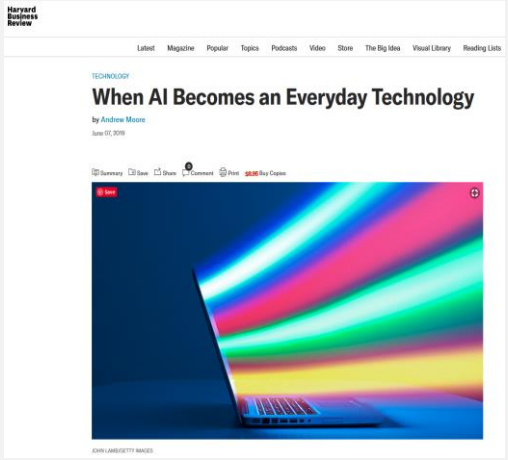
We employ machine learning techniques to identify common characteristics and features from cases in the US courts of appeals that contribute in determining dissent. Show that our models were able to predict vote alignment with an average F1 score of 73%. Exploration into which factors help in arriving at this accuracy show that the length of the opinion, the number of citations in the opinion, and voting valence, are all key factors. These results indicate that certain high level characteristics of a case can be used to predict dissent. We also explore the influence of dissent using seating patterns of judges, and our results show that raw counts of how often two judges sit together plays a role in dissent. In addition to the dissents, we analyze the notion of memetic phrases occurring in opinions - phrases that see a small spark of popularity but eventually die out in usage - and try to correlate them to dissent.

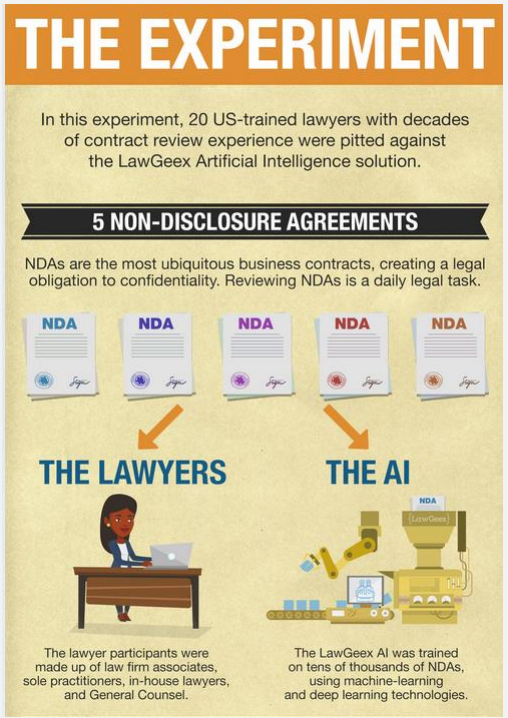
Supreme Court decisions [9], something which legal experts are notoriously unsuccessful at, or predicting authorship of unsigned judicial opinions [12].

Our overarching objective is two-fold - firstly, to predict the vote alignment between two Court of Appeals judges based on their historical voting record, as well as other case-based and judge-based features. Secondly, we consider how seating and citation patterns between judges affect their voting. Thus, in addition to using the voting history, we also make use of the citation and seating networks among judges.

2 DATA

The original dataset contains opinions from 387,898 cases (1880-2013), collected by one of the authors, as well as features for these cases from "The United States Courts of Appeals database" [14].





DE OLIVEIRA, R. B. **Utilização de Ontologias para Busca em Base de Dados de Acórdãos do STF**. 2017. 58 f. Dissertação (Mestrado) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2017.

O Supremo Tribunal Federal (STF) mantém uma base de documentos que relatam suas decisões tomadas em todos os julgamentos passados. Esses documentos são chamados de acórdãos e compõem a jurisprudência do STF, pois abordam assuntos que dizem respeito a constituição. Eles estão disponíveis a todos, porém encontrar uma informação relevante é uma tarefa árdua, que muitas vezes exige um nível de conhecimento da área jurídica.

O STF oferece um mecanismo de busca para esses acórdãos, porém o mecanismo atual utiliza uma forma tradicional de busca baseado em formulários com inúmeros campos a serem preenchidos e selecionados, se assemelhando a um questionário, no qual cada pergunta está relacionada a filtragem de certas informações em toda a base persistida em bancos de dados relacional. Esta abordagem do ponto de vista do usuário é pouco intuitiva e em alguns casos imprecisa.

Com base nesta dificuldade, neste trabalho é apresentada uma abordagem de um mecanismo de pesquisa que utiliza uma ontologia para a criação de uma representação do conhecimento contido nos acórdãos do STF. Sua construção é feita com o auxílio da tecnologia OBDA (Ontology Based Data Access), que permite a criação de uma camada semântica sobre uma base de dados relacional, o que possibilita a realização de consultas em SPARQL.

Futures Exchange Reins In Runaway Trading Algorithms

CME takes emergency measures to combat surging data volumes in Eurodollar futures

Imagem derivada de foto de Steve Jurvetson licenciada como CC-2.0-BY

Inovação

Professores do Ifes realizam chamadas por reconhecimento facial

Aplicativo de inteligência artificial desenvolvido pelo Laboratório de Extensão em Desenvolvimento de Soluções (LEDS) do Ifes da Serra facilita e reduz o tempo da chamada

Raquel Lopes
rflopes@redegazeta.com.br
Publicado em 23/10/2019 às 16h08

Considerações históricas

Considerações histórico-metodológicas

- O campo da Inteligência Artificial (IA) surgiu com o objetivo de replicar a inteligência humana de forma ampla
  - Exemplo: Visão Computacional como "projeto de verão"
  - Ciclos de entusiasmo e "invernos da IA"
- Com a dificuldade do problema, o foco mudou para a realização de tarefas específicas
  - Primeiro momento: sistemas especialistas
  - Hoje: capacidades humanas ou superiores em algumas tarefas
    - Xadrez, Go…
    - Identificação de elementos em textos jurídicos

O que é inteligência artificial?

Delimitação do objeto de discussão

As aplicações que normalmente designamos como *inteligência artificial* compartilham algumas características.

Foco/*Critério*	*Como humanos*	*Racionalmente*
Pensar	Modelos cognitivos	Leis do pensamento
Agir	Jogo da imitação	Agentes inteligentes

Tabela 1: Objetivos possíveis para a Inteligência Artificial.
Adaptado de Russell & Norvig (2010)

Definições jurídicas: Brasil (e. g. PL 21/2020), Comissão Europeia (2021) — em tramitação

Do que não falaremos

Delimitação do objeto de discussão

Não discutiremos as preocupações de longo prazo com a inteligência artificial.

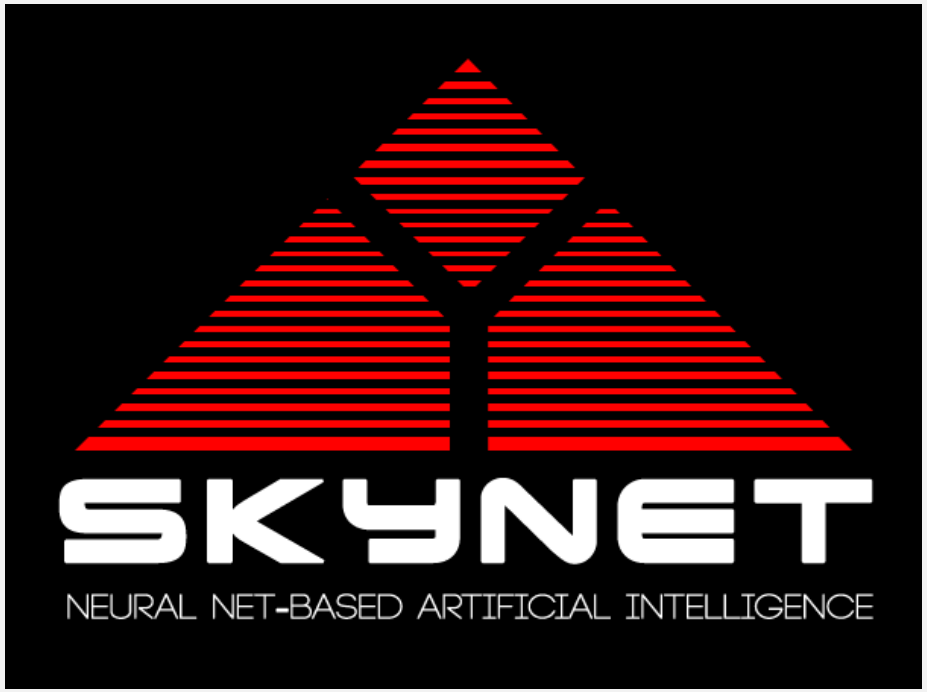
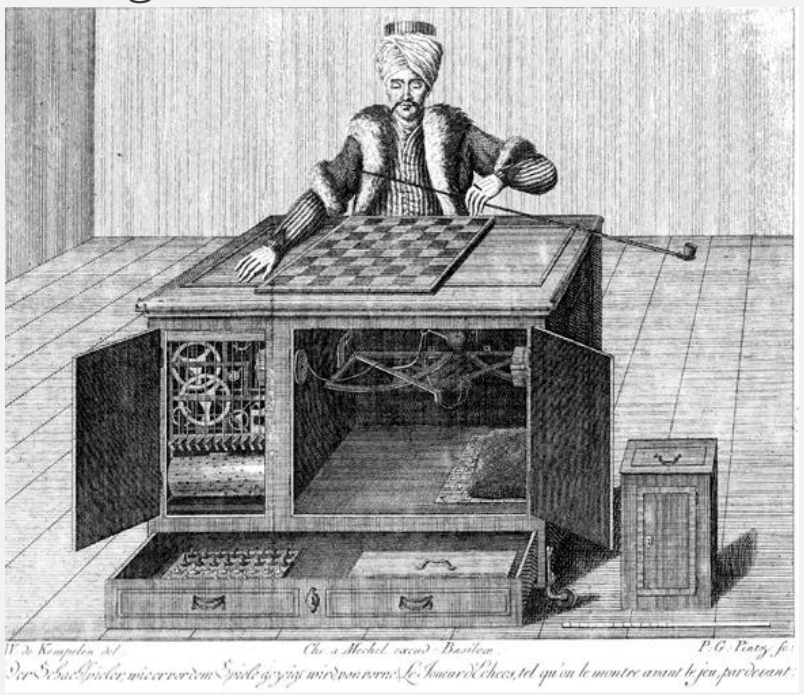


Imagem disponibilizada por Vasilisilio sob licença CC-4.0-BY-SA.

Muitas aplicações vendidas como inteligência artificial pouco têm de inteligente.



Karl Gottlieb von Windisch, 1783. *Briefe über den Schachspieler des Hrn. von Kempelen, nebst drei Kupferstichen die diese berühmte Maschine vorstellen.* Imagem em domínio público.

Inteligência artificial como questão jurídica

WORLD

It's Scarily Easy To Track Someone Around A City Via Their Instagram Stories

By cross-referencing just one hour of footage from public webcams with stories taken in Times Square, BuzzFeed News confirmed the full identities of a half dozen people.


Megha Rajagopalan
BuzzFeed News Reporter


Alison Killing
BuzzFeed Contributor


Jeremy Singer-Vine
BuzzFeed News Reporter


Hayes Brown
BuzzFeed News Reporter

Why Machine Learning May Lead to Unfairness: Evidence from Risk Assessment for Juvenile Justice in Catalonia

Songül Tolan*
Marius Miron*
Joint Research Centre,
European Commission
firstname.lastname@ec.europa.eu

Emilia Gómez
Joint Research Centre,
European Commission
Universitat Pompeu Fabra, Barcelona
emilia.gomez@upf.edu

Carlos Castillo
Universitat Pompeu Fabra, Barcelona
chato@acm.org

ABSTRACT

In this paper we study the limitations of Machine Learning (ML) algorithms for predicting juvenile recidivism. Particularly, we are interested in analyzing the trade-off between predictive performance and fairness. To that extent, we evaluate fairness of ML models in conjunction with SAVRY, a structured professional risk assessment framework, on a novel dataset originated in Catalonia. In terms of accuracy on the prediction of recidivism, the ML models slightly outperform SAVRY; the results improve with more data or more features available for training (AUCROC of 0.64 with SAVRY vs. AUCROC of 0.71 with ML models). However, across three fairness metrics used in other studies, we find that SAVRY is in general fair, while the ML models tend to discriminate against male defendants, foreigners, or people of specific national groups. For instance, foreigners who did not recidivate are almost twice as likely to be wrongly classified as high risk by ML models than Spanish nationals. Finally, we discuss potential sources of this unfairness and provide explanations for them, by combining ML interpretability techniques with a thorough data analysis. Our findings provide an explanation for why ML techniques lead to unfairness in data-driven risk assessment, even when protected attributes are not used in training.

CCS CONCEPTS

• Information systems → Expert systems; • Applied computing → Law; • Social and professional topics;

QC, Canada. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3322640.3326705

1 INTRODUCTION

Machine learning (ML) systems detect patterns in data and are able to predict complex outputs under high uncertainty [37]. Medicine, finance and law, are a few domains where humans rely on an algorithms to solve expert tasks [28]. In these cases ML systems can surpass human capabilities, particularly when dealing with large datasets or a high number of input features. One example where ML algorithms and expert systems can better inform human decisions is predicting criminal recidivism [26], defined as the act of a person committing a crime after they have been convicted of an earlier crime [11]. However, the adoption of ML in this area is problematic, knowing that the decisions of ML models can often be biased and discriminate against certain minority groups or populations, becoming unfair [3, 4, 13]. This can be concerning for nontransparent models, such as deep neural networks, if the decision process is not made transparent. [29].

Here we propose a methodology to assess predictive performance and unfairness for the ML methods used in juvenile recidivism prediction, and to investigate the potential sources of unfairness. To that extent, we compare the ML models with an existing risk assessment tool in terms of predictive performance and fairness and we use ML interpretability [29] combined with a thorough data analysis to find explanations of disparity.

CARNEGIE ENDOWMENT FOR INTERNATIONAL PEACE

RESEARCH EXPERTS PROGRAMS EVENTS TOPICS

The Global Expansion of AI Surveillance

STEVEN FELDSTEIN

SEPTEMBER 17, 2019
PAPER

A growing number of states are deploying advanced AI surveillance tools to monitor, track, and surveil citizens. Carnegie's new index explores how different countries are going about this.

FULL TEXT
FULL AI GLOBAL SURVEILLANCE INDEX

De acordo com pesquisa publicada na *Foreign Affairs*, humanos costumam a ser convencidos por textos produzidos a partir do GPT-2: em um dos grupos da pesquisa, 72% das pessoas expostas a textos gerados pelo modelo os julgaram confiáveis, sendo que, nesse mesmo grupo, 83% das pessoas expostas a textos do *New York Times* julgaram que estes eram confiáveis.

Em paralelo, um trabalho conjunto de pesquisadores do *Allen Institute for Artificial Intelligence* (AI2) e da Universidade de Washington, utilizando o modelo *GROVER*, conseguiu produzir notícias falsas mais convincentes que as produzidas por humanos.

CURRENT ISSUE
November/December 2019

FOREIGN AFFAIRS

Magazine Regions Topics Collections Book Reviews More

Killer Apps

The Real Dangers of an AI Arms Race

By Paul Scharre May/June 2019

A construção de um sistema inteligente

Etapas e questões relevantes

Definição do problema

- Estipulação dos objetivos
  - Classificação
    - Binária
    - Multiclasse
  - Regressão
- Proposição das questões a responder
  - Objetivo abstrato dá origem a uma decisão sobre o que um sistema deve prever
  - Esta decisão se reflete na escolha de variáveis de saída
- Fatores em jogo
  - Limites técnicos e práticos
  - Conhecimento disponível
  - Efeitos nocivos em potencial

Modelagem do problema

- Escolha do modelo relevante
  - Classes de algoritmos para obter valores das variáveis de saída escolhidas
  - Escolha do modelo adequado dependerá do problema a ser resolvido pelo sistema
- Treinamento do modelo escolhido
  - Ajuste dos dados ao modelo
  - Processo envolve uso de bibliotecas e ferramentas já existentes
- Algoritmo, que é um objeto lógico, passa para um sistema computacional real
  - Novos modos de falha
  - Escolha de infraestrutura: e. g., nuvem
  - Limitação de recursos afeta treinamento

Sistemas inteligentes em produção

Desafios na operação de sistema inteligentes

- Pôr um sistema em produção é trazê-lo para o mundo real
  - Resultados são utilizados para gerar efeitos concretos
  - Modelos precisam ser capazes de rodar em larga escala
  - Uso do modelo exige interfaces para os usuários
- Sistemas terão várias interações
  - Outros sistemas
  - Usuários e *stakeholders*
- Sistema pode sofrer mudanças
  - Correção de erros
  - Reaproveitamento de soluções
  - Fim de vida útil

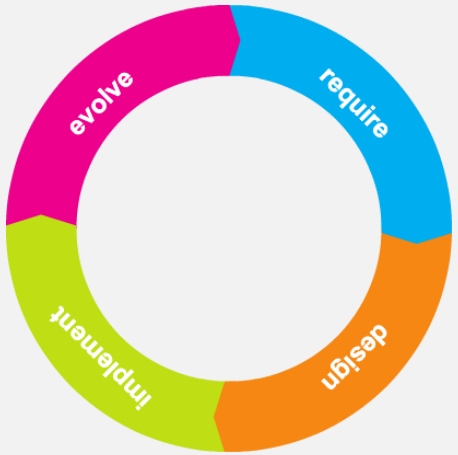


Imagem disponibilizada por Shawn Rider sob licença CC-4.0-BY-SA

Modos de falha da inteligência artificial

Como as questões juridicamente relevantes surgem?

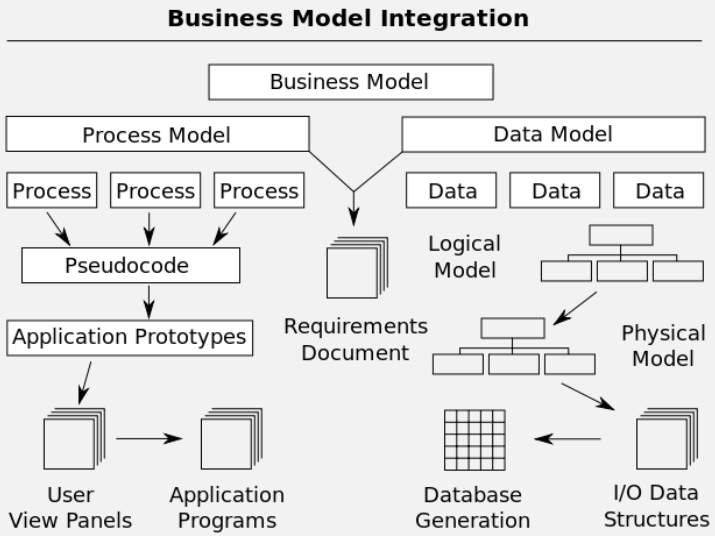
- Inteligência Artificial pode suscitar várias questões jurídicas
  - Construção dos sistemas
  - Dados utilizados
  - Escolha de modelos
  - Emprego dos sistemas construídos

- Distinguir as etapas da construção do sistema é relevante para o tratamento legal da IA
  - Imputação de responsabilidade
  - Medidas técnicas para resguardar direitos

Dados como insumo de sistemas inteligentes

Como um sistema inteligente usa dados para funcionar

- Para que um sistema inteligente possa ser usado, ele precisará de informação a respeito do contexto de uso
  - Métodos de coleta de dados
  - Qualidade dos dados obtidos
    - Representatividade
    - Consistência
  - Tratamento dos dados
    - Imputação de dados faltantes
    - *Outliers*
    - Assimetrias
  - Permissão para o uso dos dados



Dados como produto de sistemas inteligentes

Aplicações e desafios

- O uso de sistemas inteligentes gerará dados a respeito de seu objeto
  - Perfis individualizados
  - Estatísticas agregadas
  - Tomada de decisão automatizada

- Questões salientes
  - Generalização dos resultados de treinamento
  - Vieses e discriminação algorítmica
  - Explicação dos resultados produzidos
  - Possibilidade de associação entre dados e pessoas

Dados pessoais e sistemas inteligentes

Como a construção e uso da IA envolvem dados pessoais

- Muitas das aplicações que discutimos até aqui estão ligadas a pessoas naturais
  - Determinação de aspectos da personalidade
  - Previsão de comportamentos individuais e coletivos
  - Decisões que impactam a vida de pessoas
- Definições legais de dados pessoais
  - LGPD, art. 5º, I: informação relacionada a pessoa natural identificada ou identificável
  - GDPR, art. 4(1): idem, definindo "identificável" em termos da possibilidade direta ou indireta de associação a um identificador ou a um ou mais fatores específicos da identidade física, fisiológica, genética, mental, econômica, cultural ou social de pessoa natural.

Dados pessoais e sistemas inteligentes (2)

Questões salientes

- Distribuição geográfica das operações de tratamento
  - Múltiplas normas podem ser relevantes
  - Circulação dos dados gerados por sistemas inteligentes

- Muitas vezes IA é usada em contextos de *big* data
  - Dados pessoais de muitas pessoas
  - O mito dos dados anônimos: volume pode levar à identificação

- Uso de sistemas inteligentes envolve trabalho humano
  - Coleta e tratamento de dados
  - Interpretação dos resultados

Identificação com base em dados

Estimating the success of re-identifications in incomplete datasets using generative models

Luc Rocher, Julien M. Hendrickx & Yves-Alexandre de Montjoye

Nature Communications **10**, Article number: 3069 (2019) | Cite this article

83k Accesses | **2** Citations | **1861** Altmetric | Metrics

Abstract

While rich medical, behavioral, and socio-demographic data are key to modern data-driven research, their collection and use raise legitimate privacy concerns. Anonymizing datasets through de-identification and sampling before sharing them has been the main tool used to address those concerns. We here propose a generative copula-based method that can accurately estimate the likelihood of a specific person to be correctly re-identified, even in a heavily incomplete dataset. On 210 populations, our method obtains AUC scores for predicting individual uniqueness ranging from 0.84 to 0.97, with low false-discovery rate. Using our model, we find that 99.98% of Americans would be correctly re-identified in any dataset using 15 demographic attributes. Our results suggest that even heavily sampled anonymized datasets are unlikely to satisfy the modern standards for anonymization set forth by GDPR and seriously challenge the technical and legal adequacy of the de-identification release-and-forget model.

Decisões automatizadas

Contornos iniciais para discussão

- Automação de decisões pode ser vantajosa de um ponto de vista
  - Quantitativo: velocidade, custos
  - Qualitativo: capacidades, previsibilidade

- Grau de automação de decisões
  - Parcial: *Human in the loop*
  - Total: fluxos sem atuação humana
    - *Human on the loop*: casos excepcionais
    - *Human out of the loop*: autonomia
    - *Human under the loop*: restrição da autonomia humana

Exemplos de sistemas

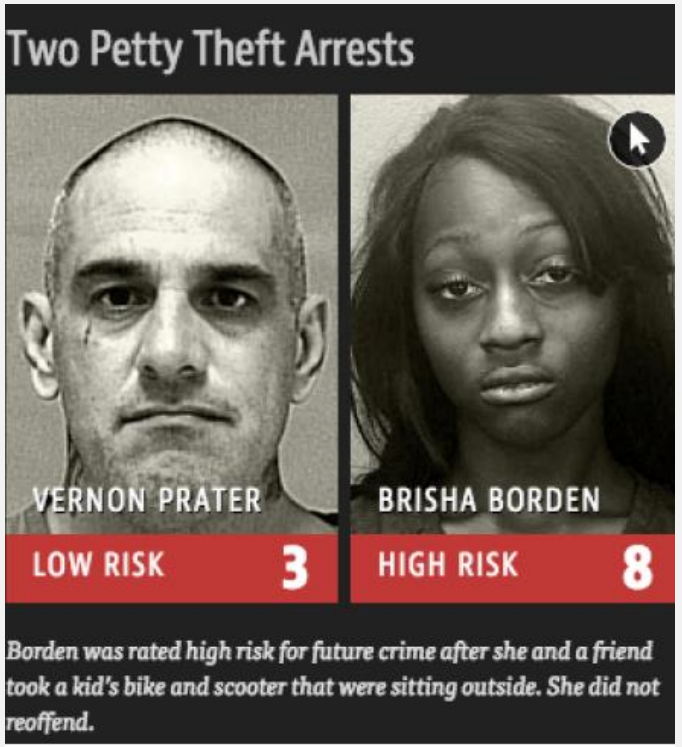
- Sistemas inteligente para triagem de currículos para uma vaga de emprego
- Sistema automatizado para avaliação de crédito e conceder ou não um empréstimo
- Sistema que indique casos julgados por um dado juiz
- Veículo não-tripulado para entregas
- Recomendação de produtos

Problemas com as decisões automatizadas

Fontes em potencial de questões jurídicas

- Questões de sistema
  - Problemas de execução
  - Problemas de projeto
  - Problemas com a finalidade

- Vieses e discriminação algorítmica
  - Viés: erro sistemático em favor de uma categoria
  - Discriminação direta e indireta
    - Direta: características protegidas
    - Indireta: escolha de *proxies*
  - Termos usados de formas distintas na computação e no Direito
  - Qual é a fonte destes problemas?



Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, "Machine Bias", 2016. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Intervenção em decisões automatizadas

Lei Geral de Proteção de Dados (BR)

Art. 20. O titular dos dados tem direito a solicitar a *revisão* de decisões tomadas *unicamente com base* em tratamento automatizado de dados pessoais que *afetem seus interesses*, incluídas as decisões destinadas a definir o seu perfil pessoal, profissional, de consumo e de crédito ou os aspectos de sua personalidade. (Redação dada pela Lei nº 13.853, de 2019)

§ 1º O controlador deverá fornecer, sempre que solicitadas, *informações claras e adequadas* a respeito dos *critérios e dos procedimentos* utilizados para a decisão automatizada, observados os segredos comercial e industrial.

§ 2º Em caso de não oferecimento de informações de que trata o § 1º deste artigo baseado na observância de segredo comercial e industrial, a autoridade nacional poderá realizar *auditoria* para verificação de aspectos discriminatórios em tratamento automatizado de dados pessoais.

§ 3º (VETADO). (Incluído pela Lei nº 13.853, de 2019)

Regulamento Geral Prot. Dados (EU)

Artigo 22.º

1. O titular dos dados tem o *direito de não ficar sujeito* a nenhuma decisão tomada *exclusivamente com base* no tratamento automatizado, incluindo a definição de perfis, que produza efeitos na sua esfera jurídica ou que o afete significativamente de forma similar.

2. O n.º 1 não se aplica se a decisão:
   a) For necessária para a celebração ou a execução de um contrato entre o titular dos dados e um responsável pelo tratamento;
   b) For autorizada pelo direito da União ou do Estado-Membro a que o responsável pelo tratamento estiver sujeito, e na qual estejam igualmente previstas medidas adequadas para salvaguardar os direitos e liberdades e os legítimos interesses do titular dos dados; ou
   c) For baseada no consentimento explícito do titular dos dados.

3. Nos casos a que se referem o n.º 2, alíneas a) e c), o responsável pelo tratamento aplica *medidas adequadas para salvaguardar os direitos e liberdades e legítimos interesses* do titular dos dados, designadamente o direito de, pelo menos, obter intervenção humana por parte do responsável, manifestar o seu ponto de vista e *contestar a decisão.* [...]

Regulação de decisões automatizadas

Questões para a efetividade da revisão

- Direito à revisão é resposta individual a lesões ou ameaças
  - Exercício deste direito exige informações sobre o que aconteceu e o que pode ser feito: *explicação*
  - Uma vez acionado o direito, a revisão precisa surtir efeitos
  - Revisão como parte de um sistema de instrumentos protetivos

- Sistemas inteligentes são artefatos complexos
  - Complexidade pode tornar opaco a construção e o uso da IA
  - *Opacidade* (Burrell 2016)
    - Matemática
    - Escala do problema
    - Contexto institucional

Como trazer a IA para a realidade notarial e registral?

- **Tratamento de dados**
  - Pré-condição para a automação
  - Cadastros únicos e fontes centralizadas
  - Separação informacional de poderes
- **Automação simples**
  - Aplicação mecânica de regras
  - Processos replicáveis
- **Automação avançada**
  - Complexidade tecnológica exige coordenação de esforços
  - Tarefas como processamento de documentos ou apoio à decisão