Fairness, Accountability, and Transparency

Lecture 4

Xiaofei Shi

Upcoming homework 1 and discussion

- Homework 1 is about reading and data collection

- Please start early and finish it on time

The growing prevalence of artificial intelligence

The AI boom is happening all over the world, and it's accelerating quickly

The second annual AI Index report pulls together data and expert field's progress and acceleration

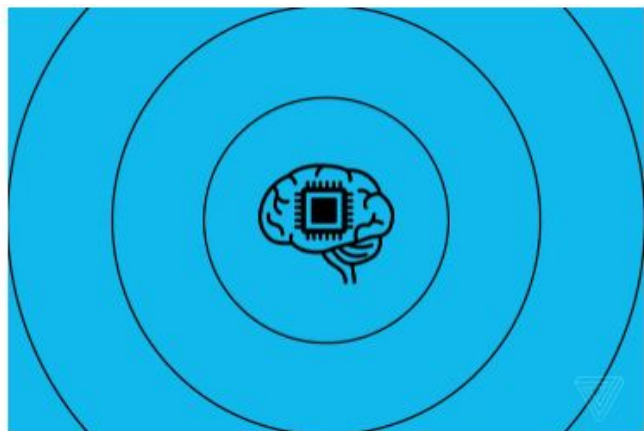By Nick Statt | @nickstatt | Dec 12, 2018, 11:00am EST

Illustration by Alex Castro / The Verge

GLOBAL MACHINE LEARNING MARKET 2020-2024

Market growth will **ACCELERATE** at a **CAGR** of almost

39%

Incremental growth ($B)

11.16

2019   2024

The market is **FRAGMENTED** with several players occupying the market

Growth Contributed by **NORTH AMERICA**

38%

Growth for **2020**

38.01%

Market impact: **POSITIVE**

17000+ Reports covering niche topics. Read them at ◀technavio

"Big data" is largely personalized data

- Examples?

"Big data" is largely personalized data

- Genetic data, medical records

- Financial history, cryptocurrency

- Browser history, online activity: amazon account, google search, social media, etc

- Sensor data, location, language, health

- .....

Learning systems increasingly adopted to filter, score, recommend, personalize, and shape human experiences

Ethical/moral/social/policy implications

- Codify and entrenching biases?

- Reduce accountability and/or transparency?

- Hinder due process?

- Infringe on personal privacy?

Historical blind spot
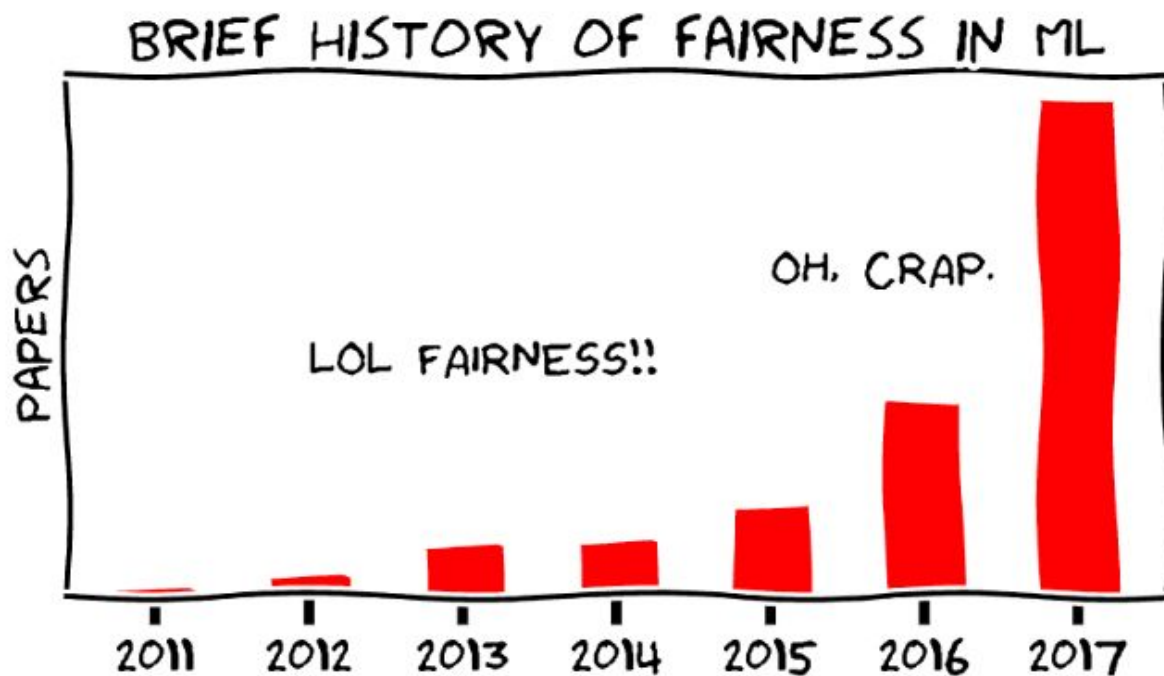
Credit: Moritz Hardt

Historical blind spot

Debate

Time
8:30pm - 9:30pm on Thursday, December 7th (Hall C). Video of debate.

Proposition
Interpretability is necessary in machine learning

Participants
Team A [*for the proposition*]: Rich Caruana (A1), Patrice Simard (A2)
Team B [*against the proposition*]: Kilian Weinberger (B1), Yann LeCun (B2)

Historical blind spot

Few computer scientists or engineers would consider developing models or tools for analyzing astronomy data without involving astronomers. So, why, then, are so many methods for analyzing social data developed without the involvement social scientists? I think, in part, it's because, as humans, we have strong intuitions

An example: OkCupid data

- In 2016, Danish researchers published data from nearly 70,000 OkCupid users, including usernames, political leanings, drug usage, and intimate sexual details

Jessie, 26
⊘ 9 miles away

First...
Car: 72 Super Beetle (bright orange with a horirble clutch)
CD: The Spice Girls 'Spice'
Kiss: Jacob Allen in my friend Krissy's basement. So many braces... so so bad.
Job: Polishing candles and firguines at The Hallmark Store.

Rachel, 29
⊘ 8 miles away

Last...
Movie I saw: Isle of Dogs
Book I read: Zadie Smith's essay collection Changing My Mind
Food I ate: A burrito
Trail I hiked: Garrapata in Big Sur
Country I visited: Ireland
Guy I swiped right on: We'll have to see now, won't we?

Malorie, 28
⊘ 6 miles away

I've never...
Driven cross country.
Had a root canal.
Missed a friend's birthday. (Really, I'm so good about birthdays. I love them.)
Met a New Pornographers fan I didn't like.

An example: OkCupid data

- In 2016, Danish researchers published data from nearly 70,000 OkCupid users, including usernames, political leanings, drug usage, and intimate sexual details

- Data released on Open Science Framework

- "Some may object to the ethics of gathering and releasing this data," the authors wrote in an associated paper, "However, all the data found in the dataset are or were already publicly available, so releasing this dataset merely presents it in a more useful form."

An example: Tinder Data

- In 2017 Stuart Colliani scraped 40,000 profile photos from Bay Area users of the Tinder data app – 20,000 from each gender

- Stuart made a facial dataset for AI experiments.

- He used facial biometric data to train a convolutional neural network

- He worked for Kaggle, a platform for machine learning and data science competitions recently acquired by Google

An example: Tinder Data

- "I have often been disappointed," Stuart writes of other facial data sets. "The datasets tend to be extremely strict in their structure, and are usually too small. Tinder gives you access to thousands of people within miles of you. Why not leverage Tinder to build a better, larger facial dataset?"

Discussion:

► Very useful and high quality data for research

► Not sure if the dating sites' term of service is updated

► Although users allowed dating sites to freely use their data, they didn't intend third parties to examine it

► Users were unable to provide informed consent

New AI can guess whether you're gay or straight from a photograph

An algorithm deduced the sexuality of people on a dating site with up to 91% accuracy, raising tricky ethical questions
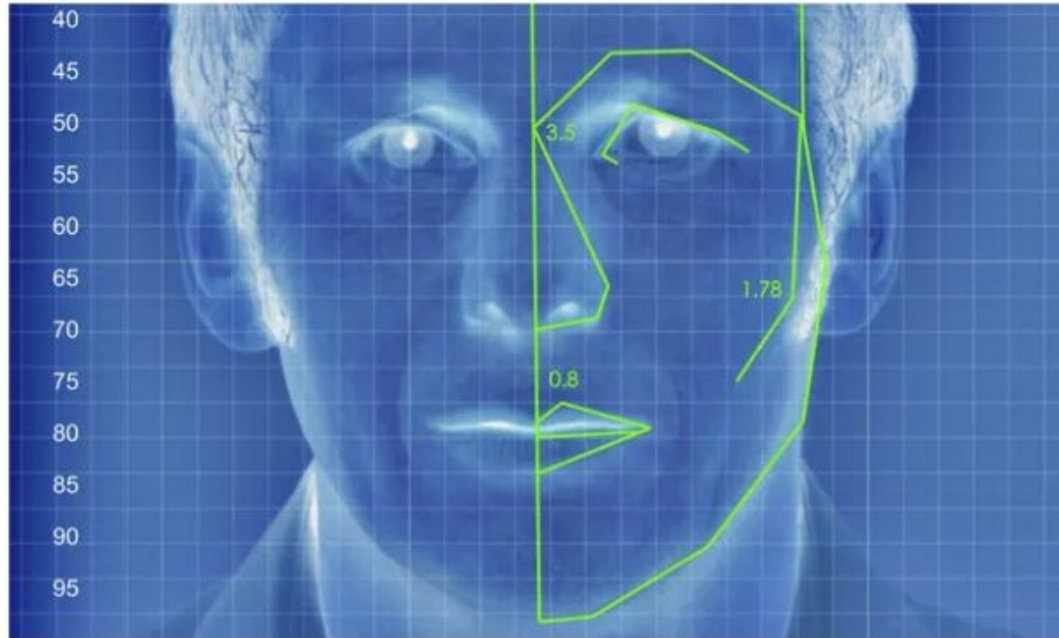
▲ An illustrated depiction of facial analysis technology similar to that used in the experiment. Illustration: Alamy

https://www.theguardian.com/technology/2017/sep/07/new-artificial-intelligence-can-tell-whether-youre-gay-or-straight-from-a-photograph

Resources

- Some useful resources:

  - [FAT* Conference](#)
  - [Wallach's blog post](#)
  - [Barocas & Hardt's NIPS17 Tutorial](#)
  - [Hardt's blog post](#)
  - [Margaret Mitchell's talk](#)
  - [Rich Caruana's talk](#)