Responsible AI
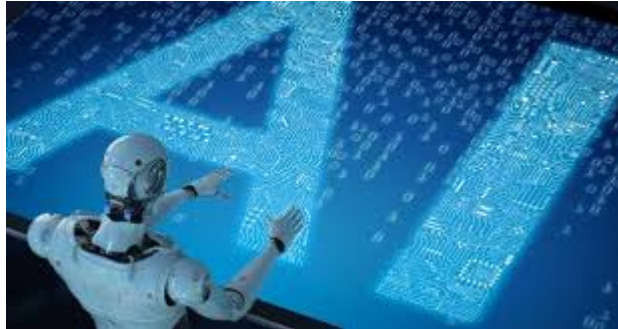
An invitation to discuss interesting problems

Hinda Haned

June 18th 2019

Ahold Delhaize

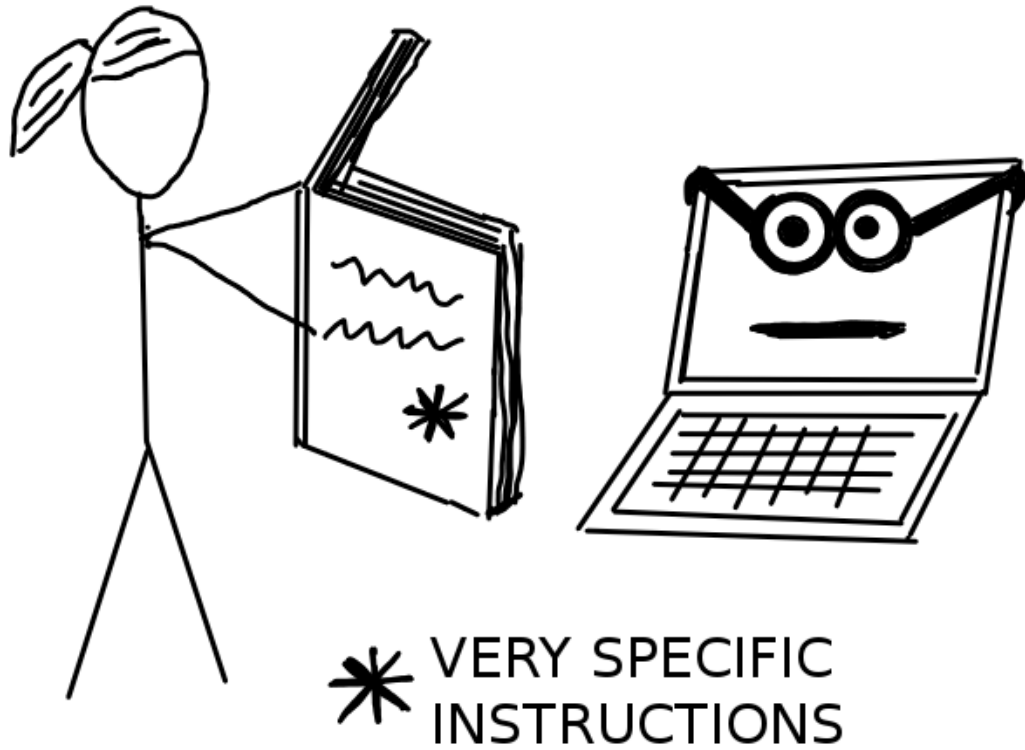University of Amsterdam

Artificial Intelligence (AI)
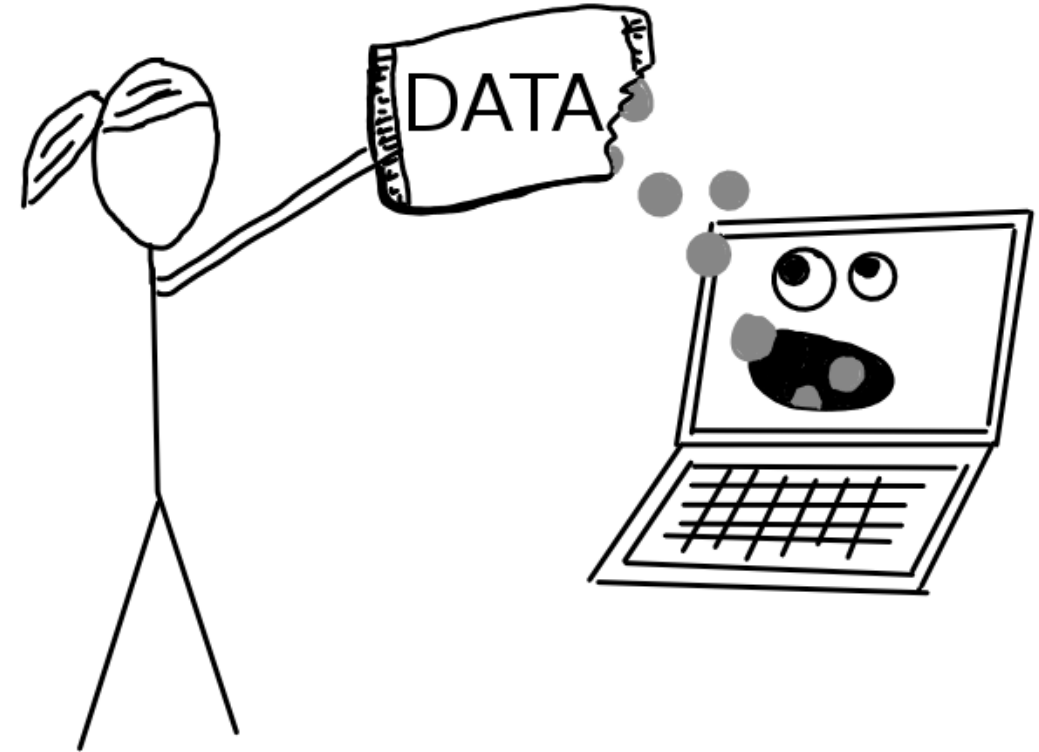
Machine Learning (ML)

Without Machine Learning

With Machine Learning

VERY SPECIFIC INSTRUCTIONS

DATA

https://christophm.github.io/interpretable-ml-book/

Machine Learning (ML)

Responsible machine learning: why bother

- 'AI hype' pushed by vendors/consultants
- GDPR 2018: 'right to an explanation' & privacy restrictions
- Critical areas in retail, e.g. associate health, recruitment, replenishment, logistics, …
- Media coverage

Controversial "AI stories"

Amazon scrapped 'sexist AI' tool
10 October 2018

Facebook's ad-serving algorithm discriminates by gender and race

Even if an advertiser is well-intentioned, the algorithm still prefers certain groups of people over others.

by Karen Hao   Apr 5, 2019

Microsoft's disastrous Tay experiment shows the hidden dangers of AI
By John West • April 2, 2016

OBSERVER

Why AI Deepfakes Should Scare the Living Bejeezus Out of You

By Harmon Leon • 06/12/19 10:36am

Bias in machine learning

$$f(\text{🗑}) = \text{🗑}$$

Data is a social mirror: perpetuating and promoting (historical) biases

Responsible ML research

- **Avoid bias** Algorithms must be designed to avoid biased outcomes
- **Fairness** Algorithms must produce fair outcomes
- **Transparency** Algorithmic outputs must be interpretable and transparent

Example: fair recruitment algorithm

- **Task** binary classification
- **Outcomes** hire/don't hire
- **Sensitive attribute** known and protected
- **Fairness goals**
  - **Group fairness** subjects in protected and unprotected groups have equal probability of being assigned to the positive prediction class
  - **Individual fairness** similar people experience similar outcomes

Fair recruitment algorithm: questions

- Group or individual fairness?
- What is the tradeoff between utility and fairness?
- What is the cost of a fair version of the current algorithm?
- How to avoid harm of representation?

Responsible ML: AI for retail Lab

A!R
ARTIFICIAL INTELLIGENCE FOR RETAIL
LAB

The **AI for Retail (AIR) LAB** conducts foundational AI research in close collaboration with academia –embedded within our brands.

UNIVERSITEIT VAN AMSTERDAM

TU**Delft**

.... | Replenishment | Search Recommend | Responsible AI | In Store Robotics | Warehouse Robotics | Delivery Robotics

In 2019 it will have 2 professors, 4 lab managers, 12 PhD's, 15 research assistants working on business relevant research topics

Responsible ML

Topic	Projects examples
Algorithmic bias discovery & mitigation	What does it mean to be fair? *e.g. introducing a taxonomy for fairness in retail applications*
Transparency in ML	ML models that provide explanations to the end-users *e.g. Why did the model make an error?*

hinda.haned@aholddelhaize.com

A!R
ARTIFICIAL INTELLIGENCE FOR RETAIL
LAB

explore & develop

The goal of our worldwide network of **AIR LABS** is to explore and develop new tech solutions, and attract, develop and retain talent. The Labs conduct foundational AI research in close collaboration with academia and provide rapid innovative tech solutions through collaborations with incubators.

Ahold Delhaize

UNIVERSITEIT VAN AMSTERDAM

TUDelft