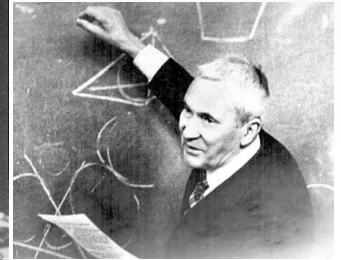
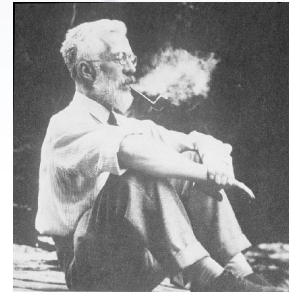
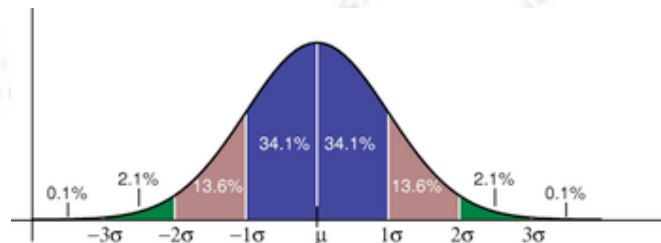


# Big Data Analysis

A discussion of Ethics in ML



Troels C. Petersen (NBI)



*“Statistics is merely a quantisation of common sense - Big Data is a sharpening of it!”*

# Ethics in Machine Learning

NOTE ON LIMITED WARRENTY!

I'm no expert on ethics, and the following are just thoughts for discussion, as I am sometimes asked about the subject!

# Ethics in Machine Learning

While ML holds many opportunities, there are certainly also some pitfalls. Many of these are of technical character, but ETHICS is also one such.

In a famous case, Target (US large supermarket chain) data mined for about 25 products, that indicated that the customer was pregnant:

One Target employee I spoke to provided a hypothetical example. Take a fictional Target shopper named Jenny Ward, who is 23, lives in Atlanta and in March bought cocoa-butter lotion, a purse large enough to double as a diaper bag, zinc and magnesium supplements and a bright blue rug. There's, say, an 87 percent chance that she's pregnant and that her delivery date is sometime in late August.

# Ethics in Machine Learning

While ML holds many opportunities, there are certainly also some pitfalls. Many of these are of technical character, but ETHICS is also one such.

In a famous case, Target (US large supermarket chain) data mined for about 25 products, that indicated that the customer was pregnant:

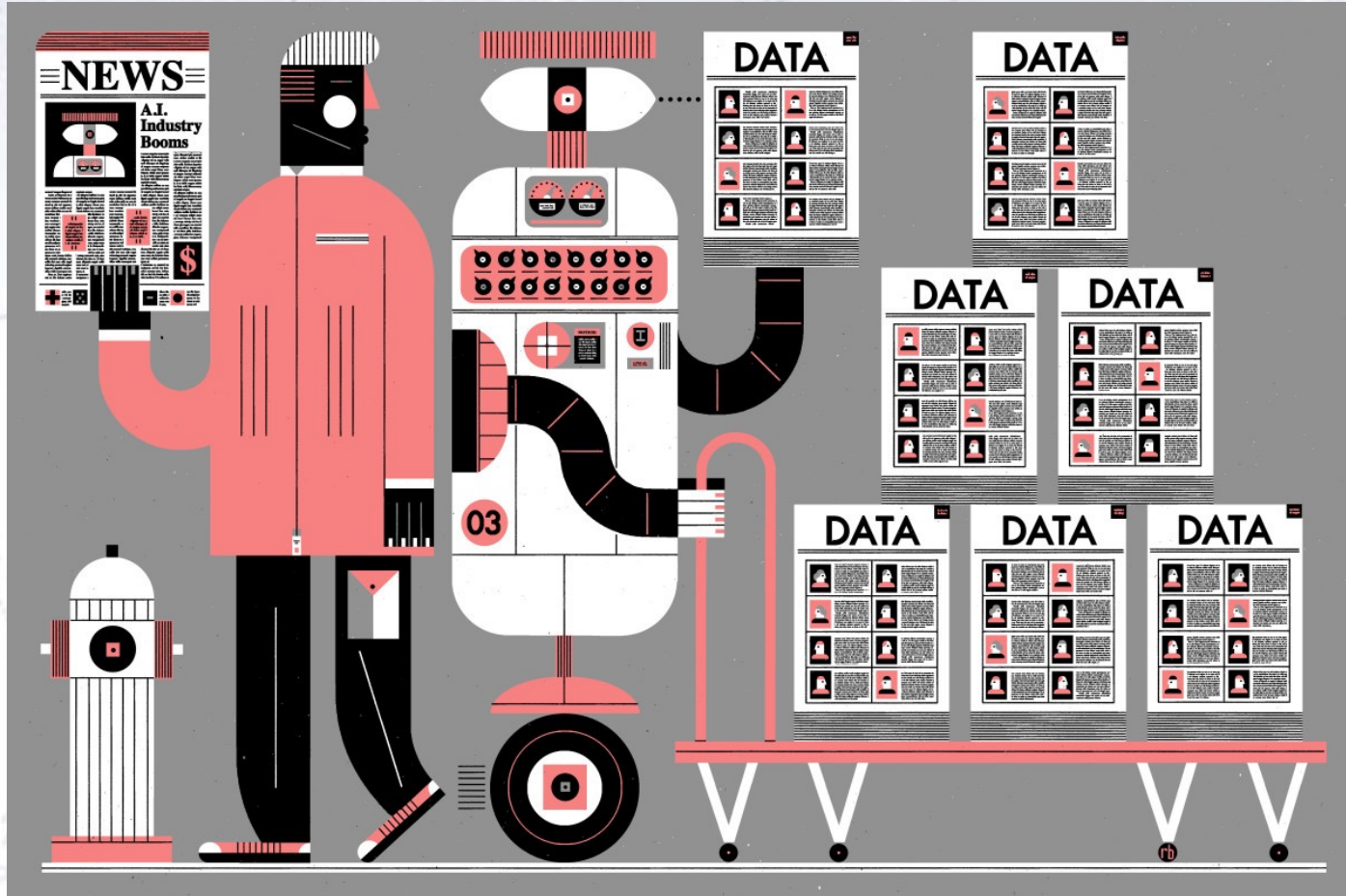
Hmmm... really? (\*)

One Target employee I spoke to provided a hypothetical example. Take a fictional Target shopper named Jenny Ward, who is 23, lives in Atlanta and in March bought cocoa-butter lotion, a purse large enough to double as a diaper bag, zinc and magnesium supplements and a bright blue rug. There's, say, an 87 percent chance that she's pregnant and that her delivery date is sometime in late August.

(\*) A link to the story in Forbes Magazine can be found by clicking this text.

# Is your software racist?

Not only may ML algorithms pick out cases “a little too well”. It may also have other “features” such as being racist!



[Link to Politico article](#)

# Is your software racist?

Not only may ML algorithms pick out cases “a little too well”. It may also have other “features” such as being racist!

**Google Translate.** Translating from Turkish, the output read like a children’s book out of the 1950’s. The ungendered Turkish sentence “o is a nurse” would become “she is a nurse,” while “o is a doctor” would become “he is a doctor.” Why? Google’s Translate tool “learns” language from existing texts, often including cultural patterns regarding how men and women are described.

**Microsoft Twitter chatbot.** It started spewing racist posts after learning from other users on the platform.

**Google’s photo-recognition.** In a particularly embarrassing example in 2015, a black computer programmer found that this tool labeled him and a friend as “gorillas.”

# Ethics in Machine Learning

It is CERTAINLY a good idea to think about the implications of using ML in ones work. Most likely it is perfectly fine (research, production, medicin, etc.), but sometimes it is less straight forward (banking, government).

I've discussed a few cases with the people involved, and here are some cases:

## Housing prices:

Would it be OK for banks and / or ministry of tax to evaluate your residence based on variables including the description from the last time it was on sale?

## Banks:

Would it be OK to ask for a meeting with a costumer, knowing that the financial distress seen in the bank is most likely because of a coming divorce?

## Schools:

Would it be OK to ask a student for a "chat", if some ML indicated that the student was about to drop out?

# Discussion

In the break-out sessions, please discuss ML ethics cases. The following might be questions that inspire (provoke?) thoughts and discussion:

- 1) In which ways do you think that ML will affect the ethics in medicine?
- 2) What changes in legislation do you think, that our the “ML world” warrants?
  - Databases
  - Collection of data
  - Surveillance
  - DNA samples
  - Browsing history
- 3) Do you think that ML will transform the way wars are fought?  
(to some extent, this has already happened).
- 4) What impact do you dream that ML will have on society, and is this match a good ethical standard (whatever that is)?