Fairness and Interpretability in Machine Learning

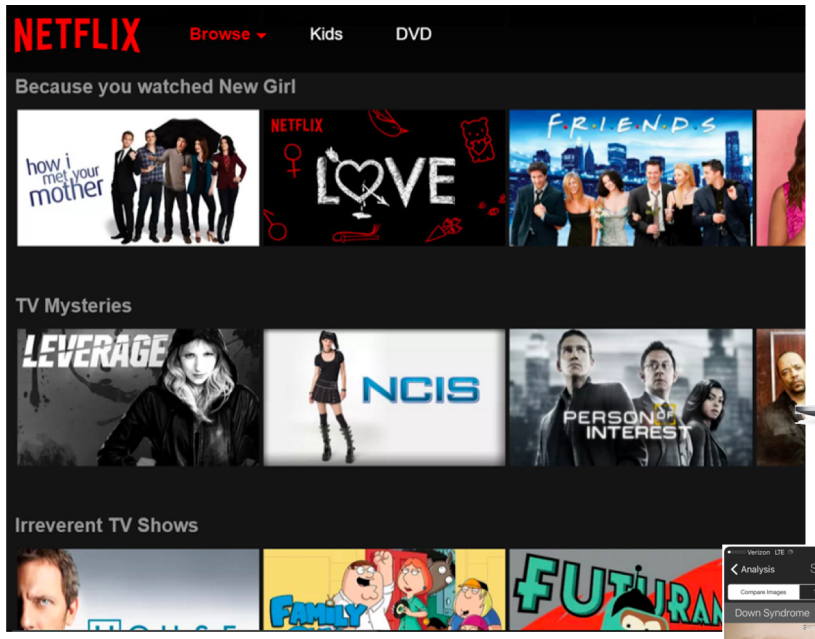
Isabel Valera

MPI for Intelligent Systems

Moscow, 2nd September 2019

ML methods are ubiquitous

[Amazon]

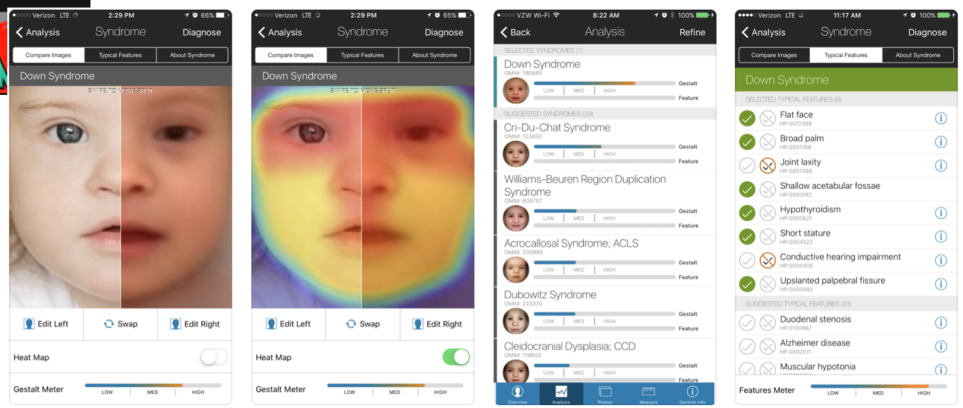


[AppleNews]

The New York Times

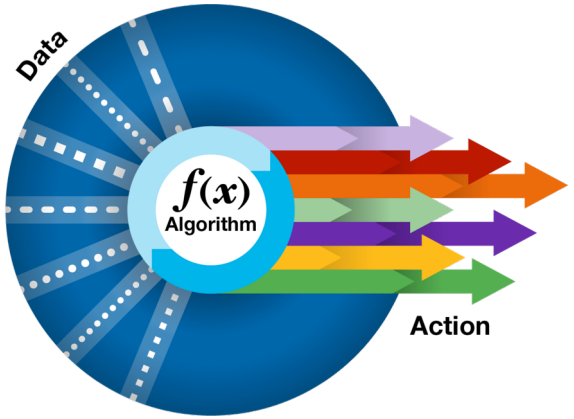
Banking Start-Ups Adopt New Tools for Lending

``Big-data lending, though, relies on software algorithms...“



[Face2Gene]

ML for decision making



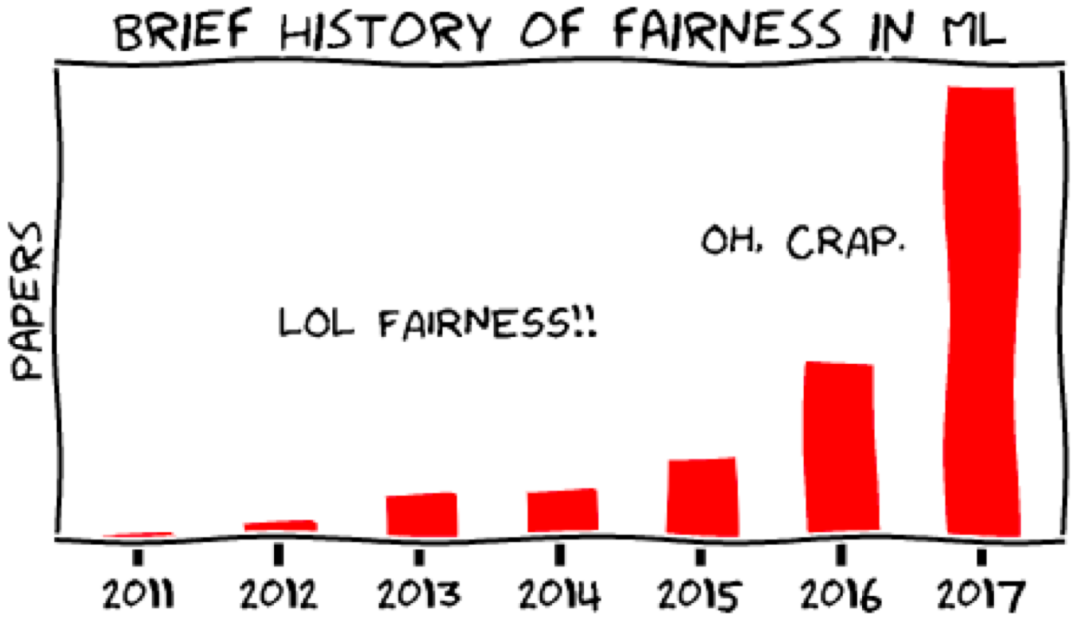
Machine learning to assist or replace human-decision making.

Increasingly used in many domains:

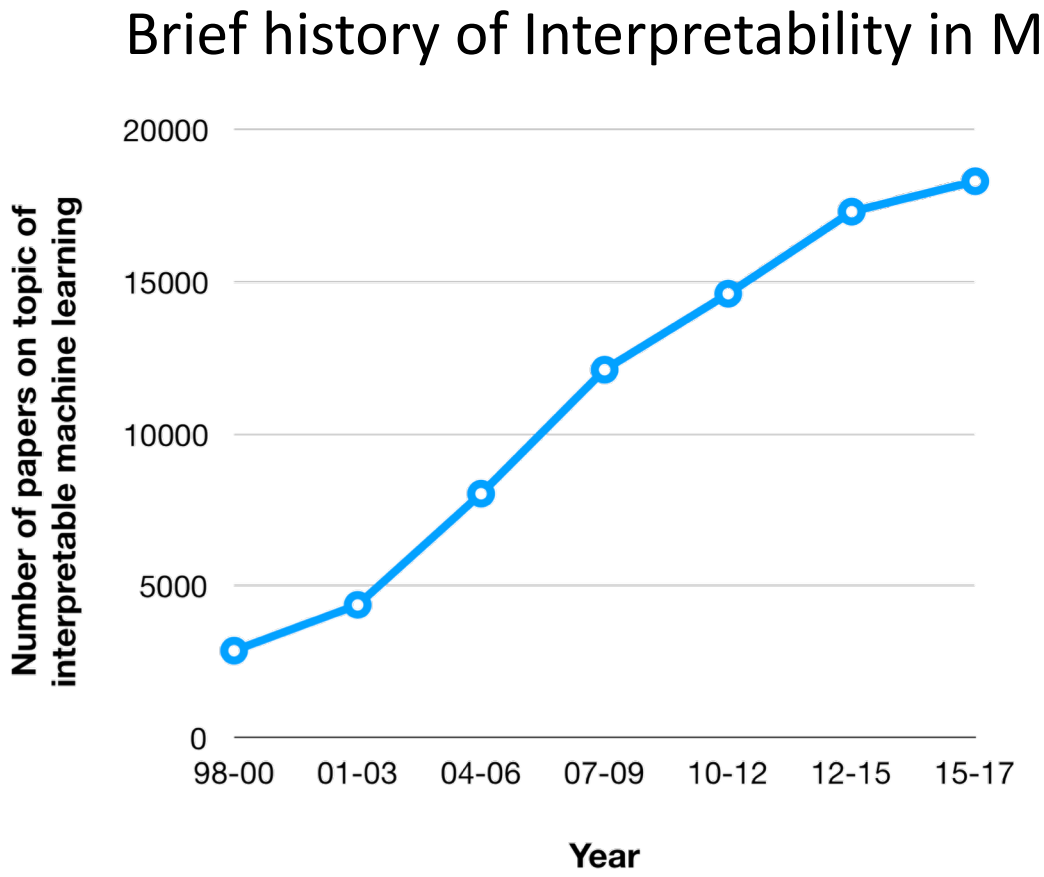
- Banking (e.g., loans)
- Insurance (e.g., premium)
- Security (e.g., patrolling)
- Reviewing (e.g., paper assignments)
- Recruiting (e.g., hiring)
- Judiciary (e.g., bail)

Social and Legal Concerns

Fairness & Interpretability in Machine Learning



[Image by Moritz Hardt]



[Image by Kim & Doshi-Velez, ICML'17]

Seminar Slides

https://github.com/probabilistic-learning/MLSS_Moscow-19

Not to forget…

Measurement and data collection processes
should be not ignored in the deployment of ML method.