



Meaningful Control of AI and Machine Ethics

Big Data in Medicine: Challenges and Opportunities

Taghi Aliyev

07/06/2019

Ethics and Sustainability

Limitations, challenges, problems

- Limited negotiation powers in decision-making
 - With Deep Learning and other recent ML-based systems
- Not all the outputs understood or explained
- Sustainability issues:
 - Need for rules and law updates constantly
- Ethical challenges:
 - Biased systems
 - Unavailability of explanations or explicit correlations

Reasonable Inference

Copyrighted Material

JUDEA PEARL
WINNER OF THE TURING AWARD
AND DANA MACKENZIE

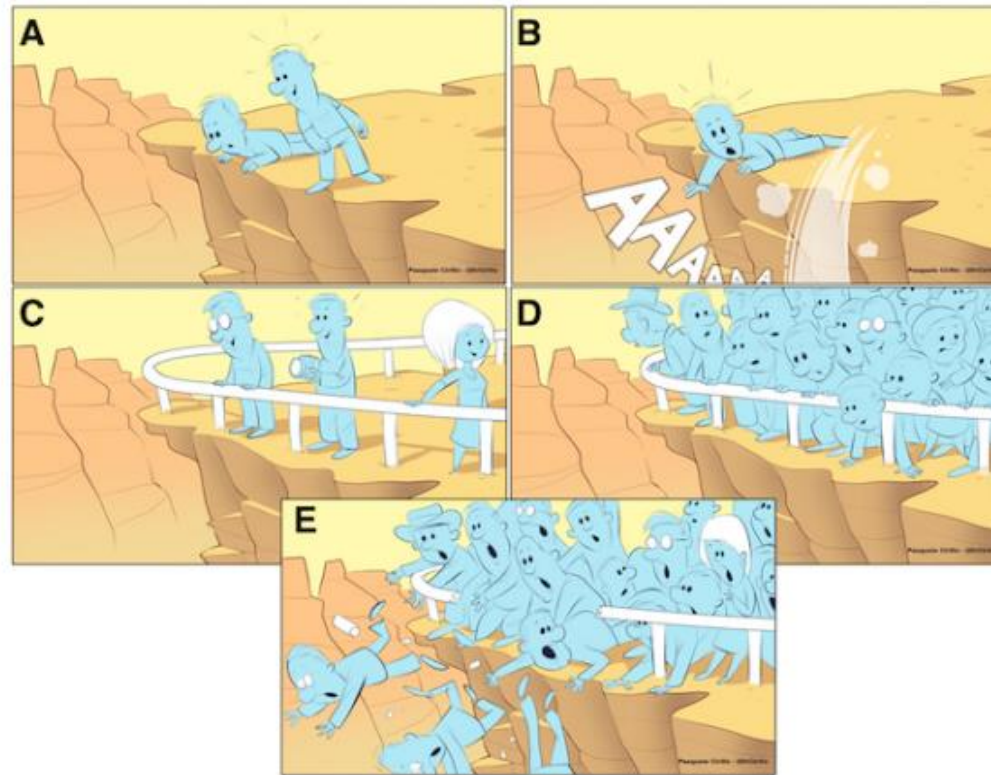
THE BOOK OF WHY



THE NEW SCIENCE
OF CAUSE AND EFFECT

Copyrighted Material

Fences and regulations



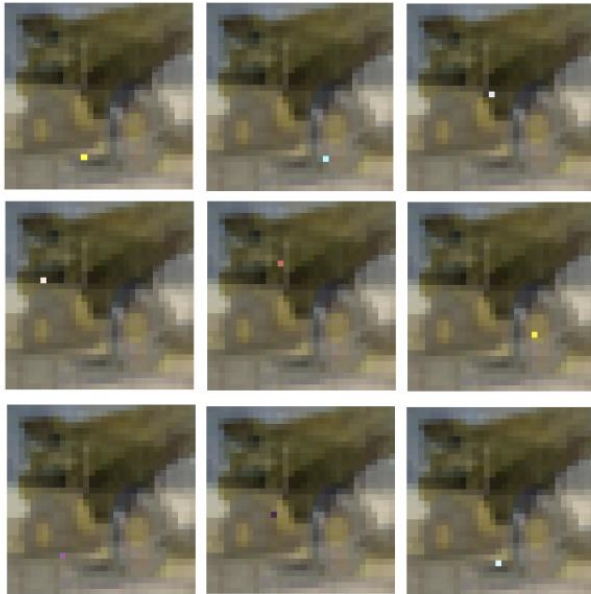
Taken from Pasquale Cirillo's "Of Risk, Fences and unavoidable falls"

Explainable AI

- Ongoing preference towards non-"black box" models
- A paradox with recent advances:
 - Better models are available, however preferred to simpler models
- Explainability of the black box models
 - Open topic
 - Next necessary step in sustainability of Deep Learning models
 - Error correction

Explainable AI

Tendencies towards non Deep Learning approaches



A White-Box Machine Learning Approach for Revealing Antibiotic Mechanisms of Action

Algorithms and Justice: Scrapping the 'Black Box'

By Cynthia Rudin | January 26, 2018



Airplane	Automobile	Bird
Cat	Deer	Frog
Horse	Ship	Truck

Target classes

Original image (dog)

Explainable AI

Required characteristics

- Negotiate the inference
- Provide useful new insight from complex modelling techniques
- Meaningful human control over the systems
- Similar with scientific findings/hypothesis:
 - If a scientist produces a new theory or finding, they need to prove it and explain it
 - Same should be upheld for ML systems

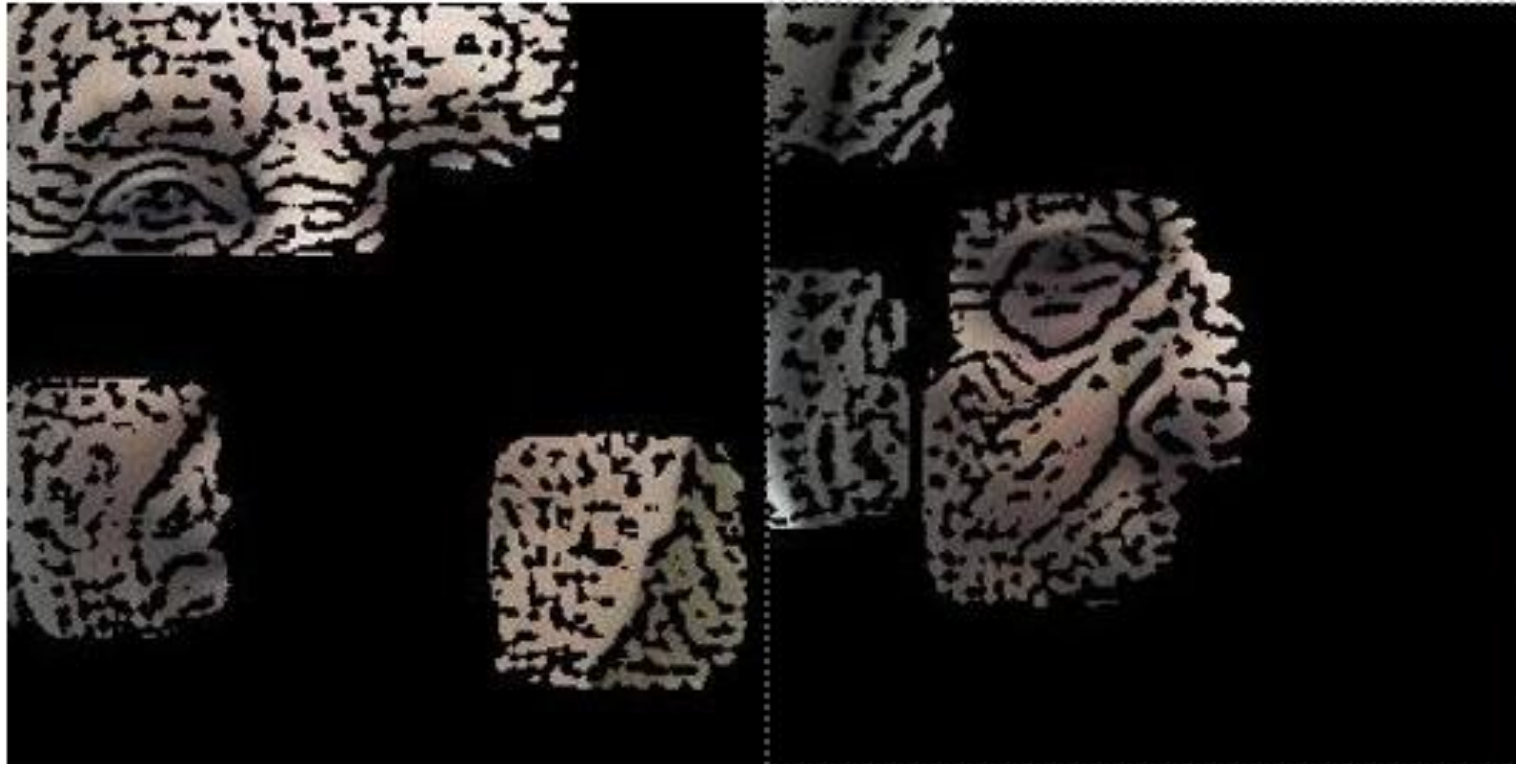
Explainable AI

Twins UK Study together with King's College London

- A step towards interpretability and meaningful human control
- Detection of facial features in twins
 - Initially heritability analysis
 - Next clinical traits and disease symptoms
- Adaptive pipeline for deconvolution
 - Based off the work from M. Zeiler (2014)

Explainable AI

Promising results



Outlook and closing remarks

- Importance of Causal Reasoning
 - Cornerstone idea for the explainable AI
- Cultural difference within AI
 - Curve Fitting/Association learning vs Casual Reasoning
- TwinsUK study an important step for explainability in Medical Research
 - Practical study and example