Explainable and convivial AI tools for healthcare

Are Francis Bacon and Ivan Illich back?

Christophe DENIS[1], Judith NICOGOSSIAN [2]

10 April 2019

[1] Associate Professor, LIP6, Sorbonne Université also EDF R&D
[2] Anthropobiologist, Neurosciences UGA
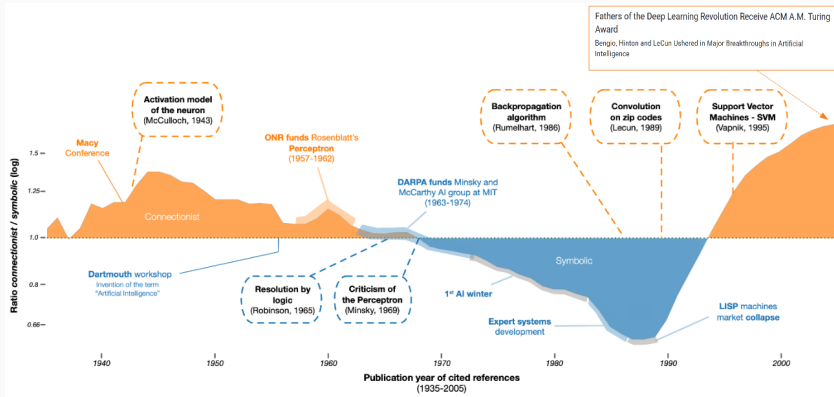contact : christophe.denis@edf.fr

Outline of the talk

Context and motivations

Our project

Focus on the evaluation of an explanation

Works in progress

AI : the revenge of neurons !

Dominique Cardon, Jean-Philippe Cointet, Antoine Mazieres. La revanche des neurones : L'invention des machines inductives et la controverse de l'intelligence artificielle. Réseaux, La Découverte, 2018
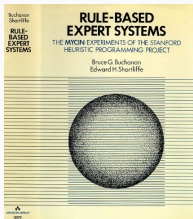
Context and motivations

Context

- Artificial Intelligence is transforming industry and in particular the medical industry.
- Towards the P4 medecine: predictive, preventive, personalized and participatory
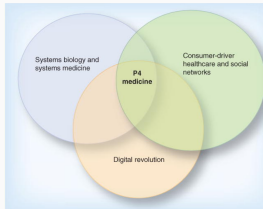
- Objectives: reducing costs while improving the healthcare system.
- Some applications: Healthcare BOTs, Virtual healthcare assistants, ...

Source: Flores, M., et al. "P4 medicine: how systems medicine will transform the healthcare sector and society".

MyCin (1970, Standford Univ.), expert system:

- identify bacteria causing severe infections, such as bacteremia and meningitis
- recommend antibiotics, with the dosage adjusted for patient's body weight

- MyCin was never actually used in practice (not for any weakness in its performance):
  1. ethical and legal issues related to the use of computers in medecine.
  2. difficulties to explain the logic of its operation and even more to detect contradictions
  3. PC's did not exist
  4. the experts had to find likelihood points for each of their inferences (unusual for them)

Today, a positive context for AI in healthcare?

- Strong need to reduce the healthcare costs
- Unlike MyCin
    - The British NHS has really used in real conditions the healthcare chatbot Babylon Health
    - First agreement from FDA to put in the market an health care expert system
- Large ecosystem dealing with AI healthcare applications

Some AI healthcare applications

Dermatoscope (2017, Standford Univ.), classification of skin cancer with deep neural networks (2,000 skin diseases)

- based on the open-source library Show and Tell (google) : a neural image caption generator
- challenge between the software and 21 dermatologists
  - dermatologists: identification of 95% of malignant tumors and 76% of the time
  - machine : identification of 96% of malignant tumors and 90% of the time

- How to validate results ? biopsy and histopathological examination
- Medical liability waiver for no dedecting malignant tumors.

Some AI healthcare applications

Decagon (2018, Standford Univ.), Modeling polypharmacy side effects with graph convolutional networks

- prevention of dangerous drug interactions for patients with complex diseases or concomitant.
- Clinical trials, in relative numbers, do not predict all interactions due to a high combinatorics.
- Graph Convolutional Network accurately predicts polypharmacy side effects, outperforming baselines by up to 69%.

M. Zitnik, M. Agrawal, J. Leskovec, Modeling polypharmacy side effects with graph convolutional networks, Bioinformatic, 2018

Beware of the AI dark side in healthcare !

1. Acceptation & Ethics
2. New & exceptional health cases
3. Human Factors: relationship between the doctor and the patient (empathic relation)
4. Loss of competence and of doubt capacity
5. Treatement monopol
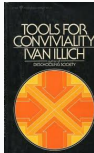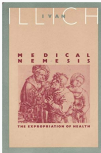
- critical analysis of the contemporary medicine examining the clinical and social reality of personal health

- argument for individuals to regain the integrity of their bodies and lives

- define the notion of the convivial tool enhancing the user's capability to work with independent efficiency (not only user friendliness !)

- e.g. MyCin was not a convivial tool, difficulties from medecine experts to define likelihood points for each of their inferences.
- A philosophical and anthropoligical approach is required to define the conviviality of an AI tool. The explainibility of the AI tools is the first step but not sufficient.

9

Our project

Our project

- Design of explainable and convivial healthcare tootls.
    - AI virtual healthcare assistants
    - Anamnesis facilitated by AI
- Need for a multi-disciplinary scientific :
    1. computer science: developpement of these tools using the appropriate technology to obtain the expected properties.
    2. computer science and epismetology: providing an explanation of the machine learning predicition (digital-symbolic coupling)
    3. At the philosophical and anthropological level, define the quality of an explanation and the humain/AI interaction.

An anthropological standpoint

- Need to integrate the tool in the flow of the consultation for the patient and in the workflow for the doctor.
- Reproduce the empathic relationship between the doctor and the patient promoting the healing
- Take into account nonverbal communication acts (emotions, gestures)
- Define the nature and the properties of the "phygital" communication, between physical and digital

Focus on the evaluation of an explanation

Is it then soon the end of model and theory?

"Petabytes allow us to say: Correlation is enough. We can stop looking for models. We can analyze the data without hypotheses about what it might show. We can throw the numbers into the biggest computing clusters the world has ever seen and let statistical algorithms find patterns where science cannot."
*Chris Anderson, The End of Theory: The Data Deluge Makes the Scientific Method Obsolete, Wired, 2008*
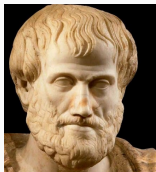
The role played by the data :

- The data is used to instantiate the problem
- In a process of validation and quantification of uncertainties, other data, moe often measurement, are used to calibrate and validate the modem

  William L. Oberkampf, Christopher J. Roy, Verification and Validation in Scientific Computing Cambridge, 2015.
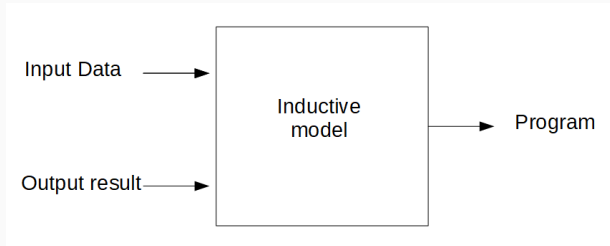
The role played by the data :

- The data are used to construct and validate the program

But inductive raisonnemment should allow from the experimental data to generalize the prediction ..

The crucial need of explainability for Machine Learning

Explainability: The ability to provide a set of users with an explanation of the results obtained by the algorithm, adapted to their scientific and business knowledge.

The need of an epistemic clarification

Machine learning weakness:

1. no causality between input and output parametres
2. the interpretability of the computational process does not by itself ensure its explicability.
3. the explanation could generate bias and be more persuasive than informative (ethical problem)
4. the preprocessing and the data selection induces ontological precursors.

Current work done with Franck Varenne.

A short taxonomy of Machine Learning explanations methods

- Linear Proxy Models: LIME, SHAP, ...
- Automatic-Rule Extraction: KT, ..
- Decision Trees: DeepRED, ..
- Salience Method: LRP , DeepLIFT, ...

Evaluation of Machine Learning explanations

An explanation of a black box can be evaluated according to its ability to

- by interpreting its mechanisms in an understandable way for the users.
- by describing precisely and exhaustively its internal mechanisms.

It is difficult to generate explanations that are both comprehensive and complete since:

1. the most accurate explanation is difficult to interpret for all users;
2. an easily interpretable explanation does not reveal the sophistication of the prediction.

Need to define a compromise between intelligibility and completeness

Works in progress

Works in progress

- Evaluation of Machine Learning explanations (collaboration with craft-ai): looking for now a trainee strudent !
- Conception of an explainable and convivial healthcare tool
- Take into account in the tool nonverbal communication acts (emotions, gestures)
- Define the nature and the properties of the "phygital" interaction, between physical and digital

That's all folks !

Thank you for your attention ! Questions ?