

L'intelligenza artificiale nel sistema delle cure: opportunità e rischi



Prof. Ing. Federico Cabitza, PhD
IRCCS Istituto Ortopedico Galeazzi
Università degli Studi di Milano-Bicocca
(Dipartimento di Informatica, Sistemistica e Comunicazione)



CONFLICT OF
INTEREST



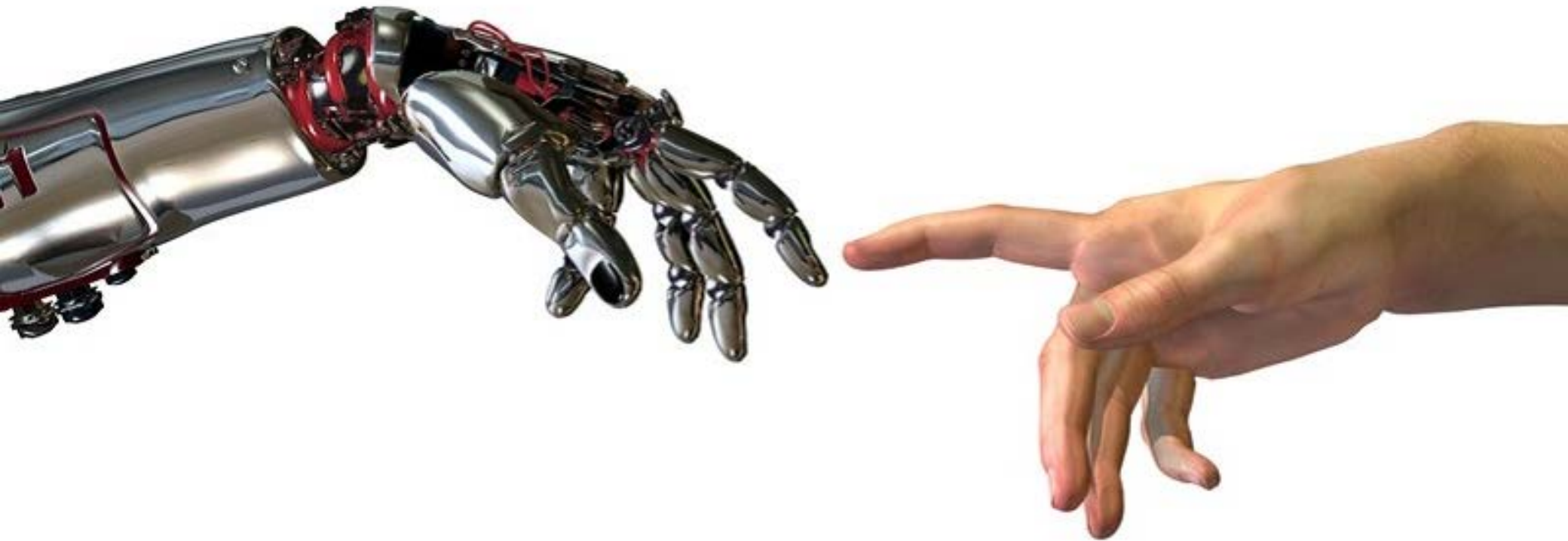
CONFLICT OF INTEREST

- ✓ Sono un ingegnere informatico (alma mater polimi);
- ✓ Sono professore di interazione uomo-macchina ad informatica;
- ✓ Insegno interazione uomo-dati (data visualization) a Data Science;
- ✓ Uso i social media per fare divulgazione su questi temi;
- ✓ Lavoro con i medici tutti i giorni...

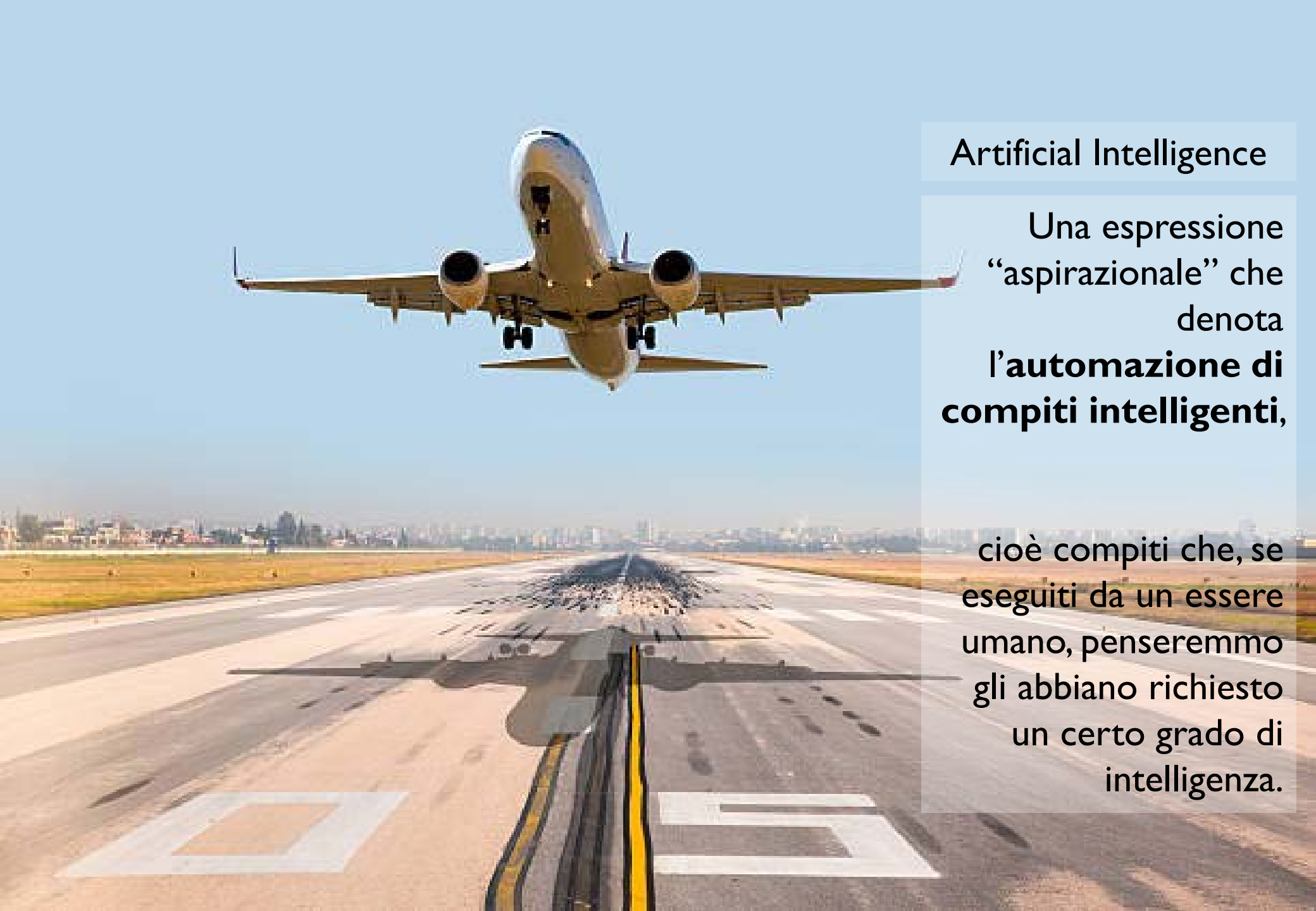




*Che tipo di
ingegnere informatico?*



Informatica come studio e progetto della **Interazione Uomo-Macchina**
Interazione Uomo-Dato (interpretazione dataset)
Interazione Uomo-AI (interpr. supporto decisionale)



Artificial Intelligence

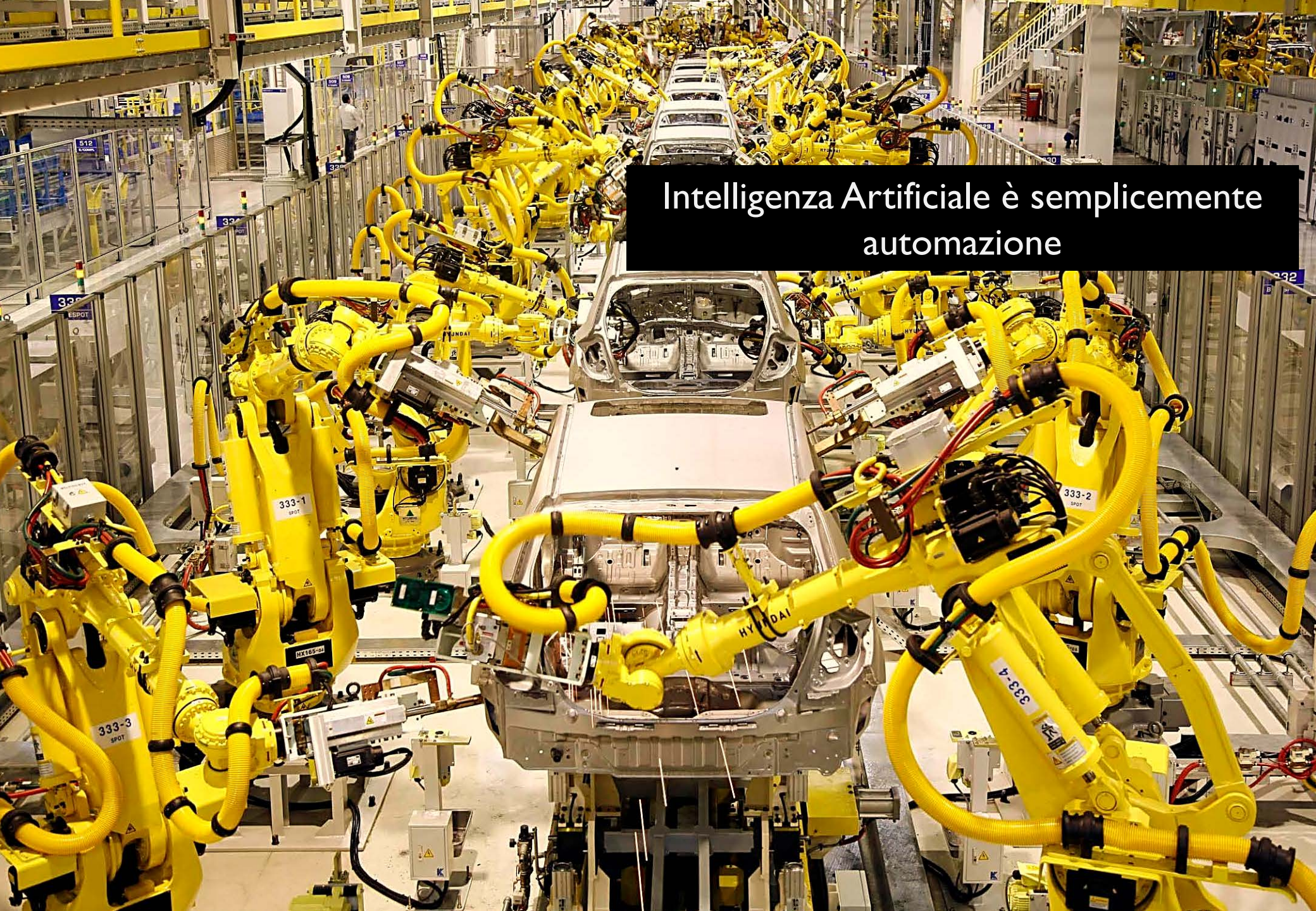
Una espressione
“aspirazionale” che
denota
**l’automazione di
compiti intelligenti,**

cioè compiti che, se
eseguiti da un essere
umano, penseremmo
gli abbiano richiesto
un certo grado di
intelligenza.

Artificial Intelligence



L'area professionale a cavallo di logica formale, ingegneria e informatica in cui operano e lavorano quelli che costruiscono sistemi AI.

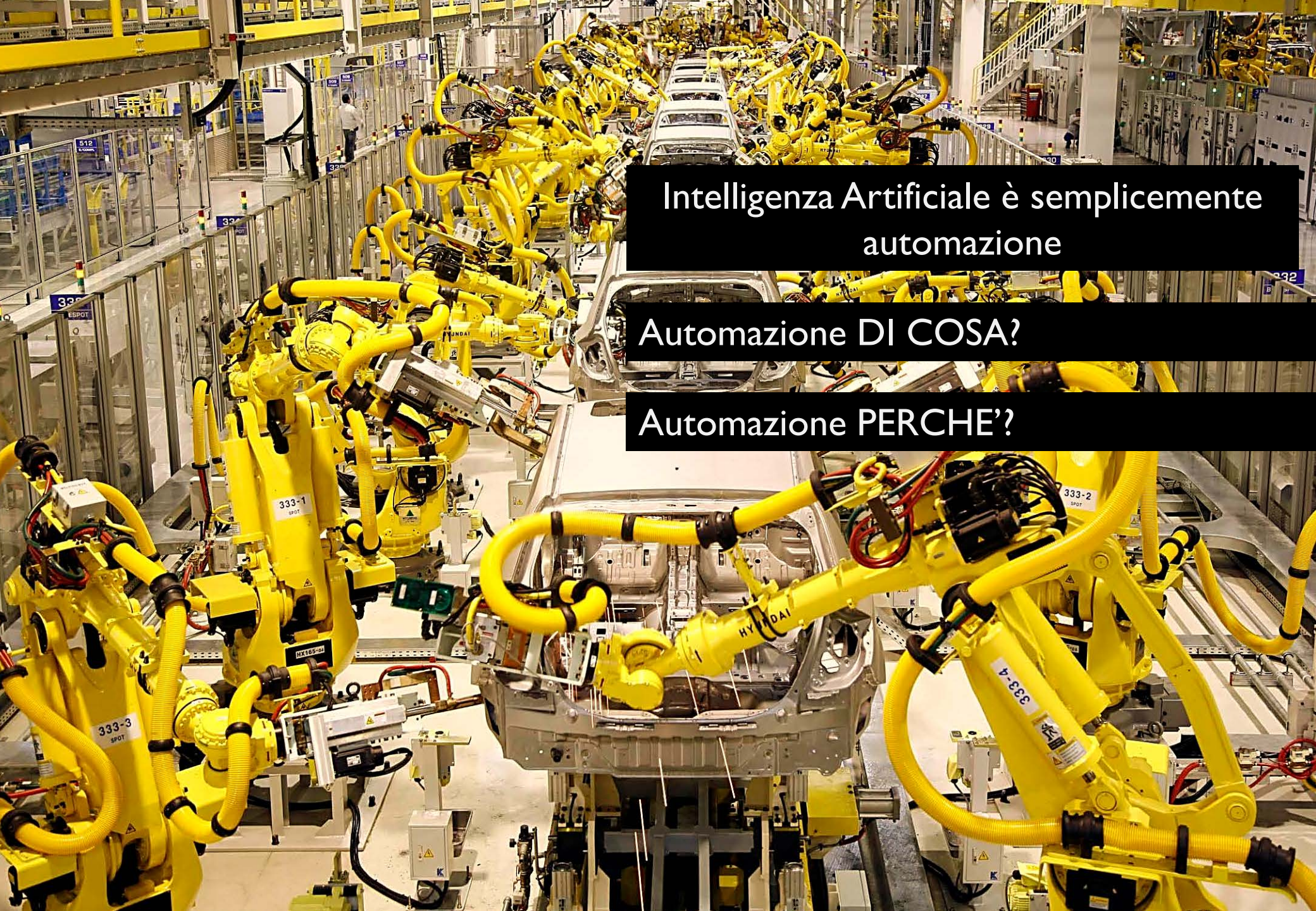


Intelligenza Artificiale è semplicemente automazione



Intelligenza Artificiale è semplicemente automazione

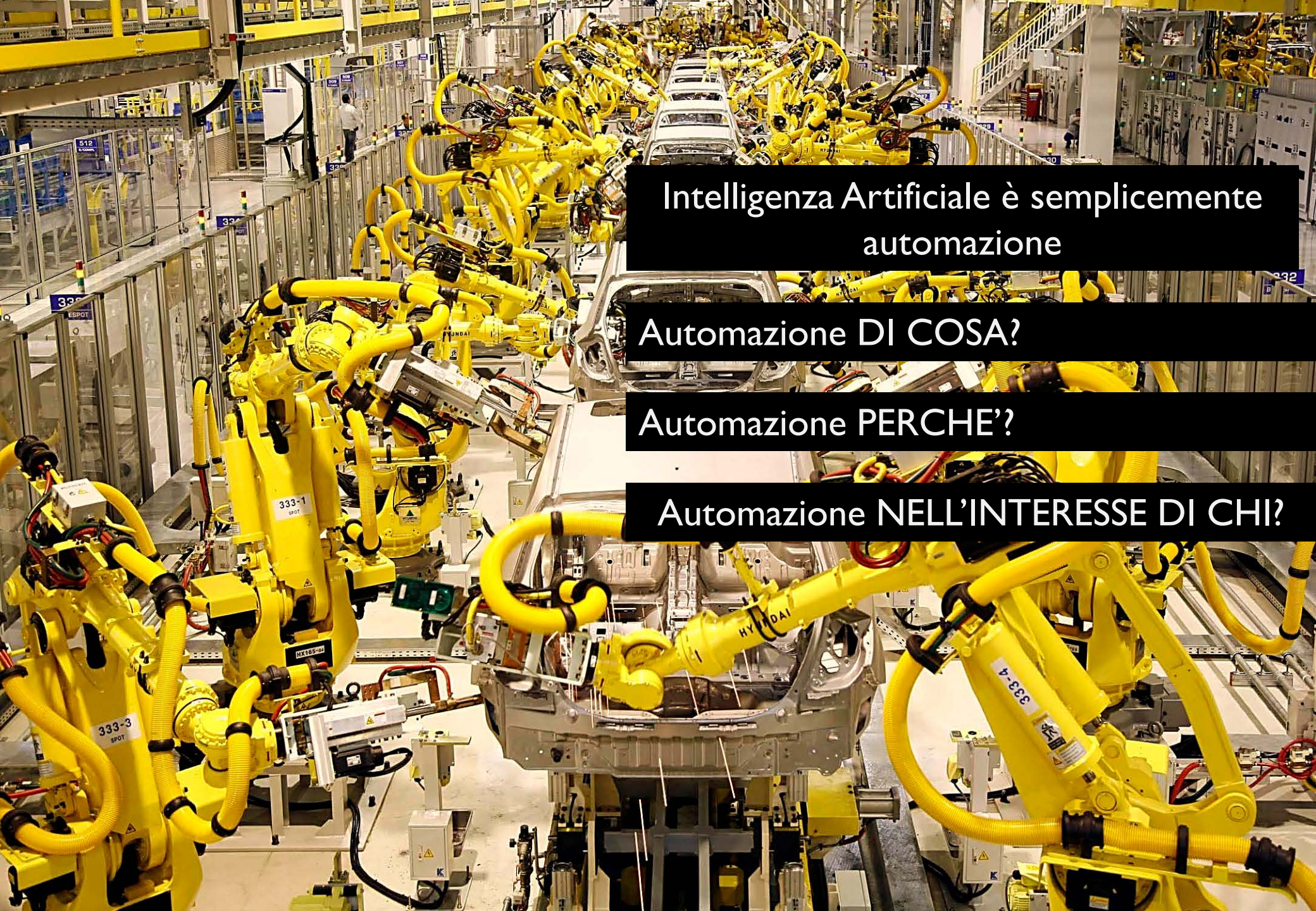
Automazione DI COSA?



Intelligenza Artificiale è semplicemente automazione

Automazione DI COSA?

Automazione PERCHE'?



Intelligenza Artificiale è semplicemente automazione

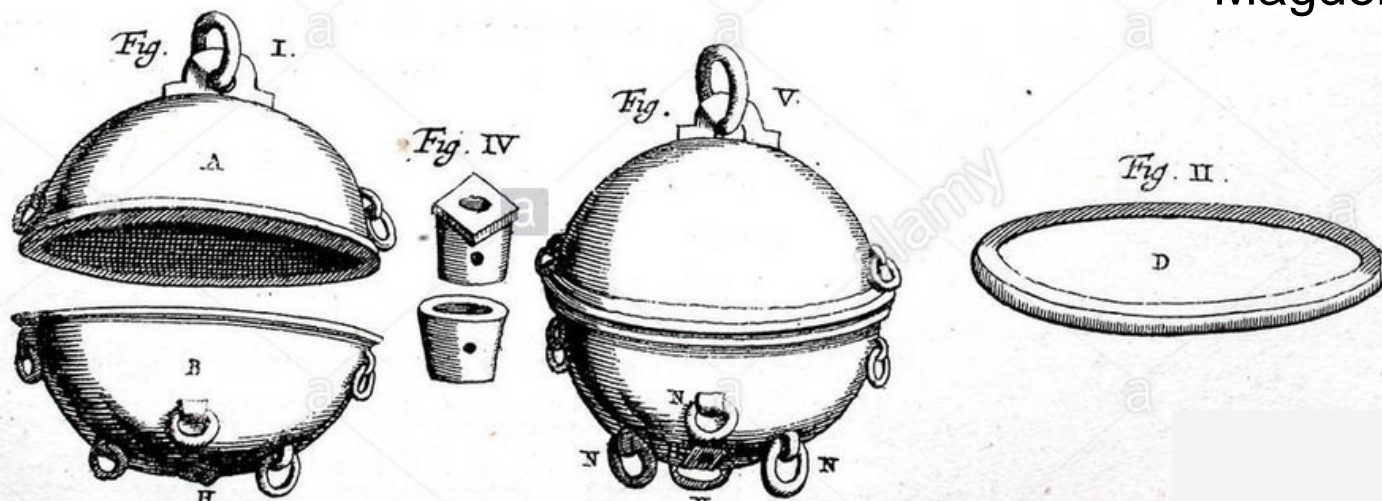
Automazione DI COSA?

Automazione PERCHE'?

Automazione NELL'INTERESSE DI CHI?

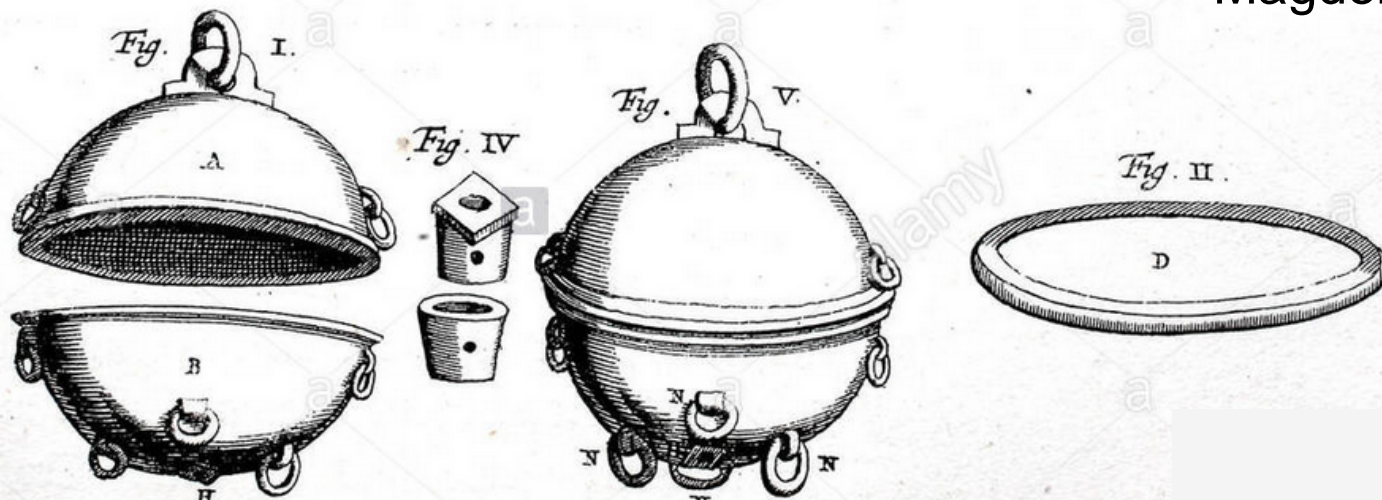


Magdeburg 1654





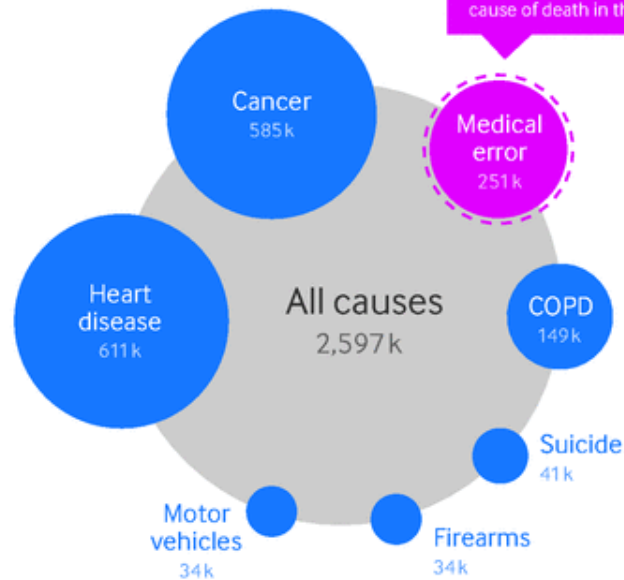
Magdeburg 1654



EFFICACIA

EFFICIENZA

Causes of death, US, 2013



Based on our estimate, medical error is the 3rd most common cause of death in the US

However, we're not even counting this - medical error is not recorded on US death certificates

© 2016 BMJ Publishing group Ltd.

Data source:
http://www.cdc.gov/nchs/data/nvsr/nvsr64/nvsr64_02.pdf



BMJ 2016;353:i2139 doi: 10.1136/bmj.i2139 (Published 3 May 2016)

Page 1 of 5



ANALYSIS

Medical error—the third leading cause of death in the US

Medical error is not included on death certificates or in rankings of cause of death. **Martin Makary** and **Michael Daniel** assess its contribution to mortality and call for better reporting

Martin A Makary *professor*, Michael Daniel *research fellow*

Department of Surgery, Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA

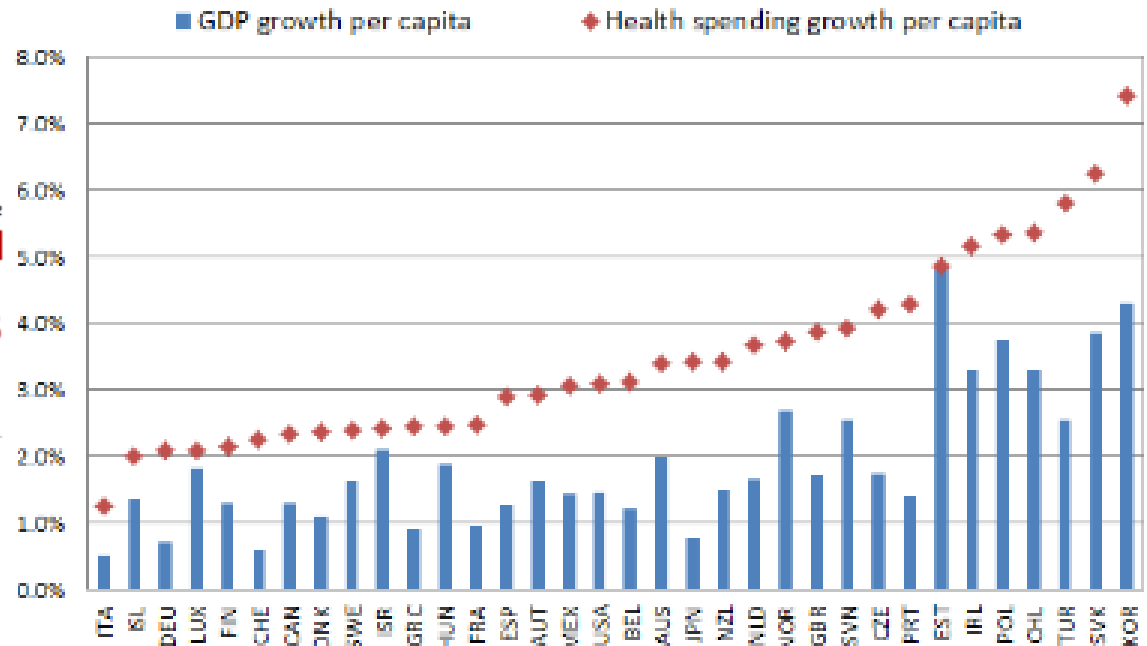
EFFICACIA

EFFICIENZA



Health spending has outpaced economic growth

Average growth rate of health spending and GDP per capita, 1990-2012



thebmj

BMJ 2017;358:j3360 doi: 10.1136/bmj.j3360 (Published 14 July 2017)

Page 1 of 2

EDITORIALS



Burnout among doctors

A system level problem requiring a system level response

Jane B Lemaire *clinical professor*¹, Jean E Wallace *professor*²

¹Cumming School of Medicine, University of Calgary, Alberta, Canada; ²Department of Sociology, Faculty of Arts, University of Calgary

Although doctors have a professional responsibility to be at the

disorganised rotations and inadequate supervision are also

10 AI Applications That Could Change Health Care

APPLICATION	POTENTIAL ANNUAL VALUE BY 2026	KEY DRIVERS FOR ADOPTION
Robot-assisted surgery	\$40B	Technological advances in robotic solutions for more types of surgery
Virtual nursing assistants	20	Increasing pressure caused by medical labor shortage
Administrative workflow	18	Easier integration with existing technology infrastructure
Fraud detection	17	Need to address increasingly complex service and payment fraud attempts
Dosage error reduction	16	Prevalence of medical errors, which leads to tangible penalties
Connected machines	14	Proliferation of connected machines/devices
Clinical trial participation	13	Patent cliff; plethora of data; outcomes-driven approach
Preliminary diagnosis	5	Interoperability/data architecture to enhance accuracy
Automated image diagnosis	3	Storage capacity; greater trust in AI technology
Cybersecurity	2	Increase in breaches; pressure to protect health data

SOURCE ACCENTURE

© HBR.ORG

https://hbr.org/2018/05/10-promising-ai-applications-in-health-care?platform=hootsuite

ANALISI E INTERPRETAZIONE DEI SEGNI, SINTOMI, IMMAGINI E DATI DI LABORATORIO PER FORMULARE DIAGNOSI E RACCOMANDAZIONI

APPLICATION

POTENTIAL ANNUAL VALUE BY 2026

Robot-assisted surgery



\$40B

Virtual nursing assistants



20

Administrative workflow



18

Fraud detection



17

Dosage error reduction



16

Connected machines



14

Clinical trial participation



13

Preliminary diagnosis



5

Automated image diagnosis



3

Cybersecurity



2

SOURCE ACCENTURE

JAMA | Original Investigation | INNOVATIONS IN HEALTH CARE DELIVERY

Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs

Varun Gulshan, PhD, Lily Peng, MD, PhD, Marc Coram, PhD, Martin C. Stumpe, PhD, Derek Wu, BS, Arunachalam Narayanaswamy, PhD, Subhashini Venugopalan, MS, Kasumi Widner, MS, Tom Madams, MEng, Jorge Cuadros, OD, PhD, Ramasamy Kim, OD, DNB, Rajiv Raman, MS, DNB, Philip C. Nelson, BS, Jessica L. Mega, MD, MPH, Dale R. Webster, PhD

IMPORTANCE Deep learning is a family of computational methods that allow an algorithm to program itself by learning from a large set of examples that demonstrate the desired behavior, removing the need to specify rules explicitly. Application of these methods to medical imaging requires further assessment and validation.

OBJECTIVE To apply deep learning to create an algorithm for automated detection of diabetic retinopathy and diabetic macular edema in retinal fundus photographs.

Editorial pages 2366 ar
Supplemental content

SCIENTIFIC REPORTS

OPEN

Discrimination of Breast Cancer with Microcalcifications on Mammography by Deep Learning

Chang Tan¹, Cangzheng Jin¹ & Li Li¹

breast cancer. To improve the diagnostic accuracy and performance of deep learning-based models on large segmentation method was used to characterize the model was constructed to assess the accuracies of the method or combination, for classifying breast lesions. Our deep learning model achieved a discriminative performance characterized alone, compared to 85.8% with a support methods with masses alone and improved to microcalcifications. Image segmentation with results for the three scenarios, respectively. Overall,

npj | Digital Medicine

www.nature.com/npjdigitalmed

ARTICLE OPEN

Scalable and accurate deep learning with electronic health records

Alvin Rajkomar^{1,2}, Eyal Oren¹, Kai Chen¹, Andrew M. Dai¹, Nissan Hajaj¹, Michaela Hardt¹, Peter J. Liu¹, Xiaobing Liu¹, Jake Marcus¹, Mimi Sun¹, Patrik Sundberg¹, Hector Yee¹, Kun Zhang¹, Yi Zhang¹, Gerardo Flores¹, Gavin E. Duggan¹, Jamie Irvine¹, Quoc Le¹, Kurt Litsch¹, Alexander Mossin¹, Justin Tansuwan¹, De Wang¹, James Wexler¹, Jimbo Wilson¹, Dana Ludwig², Samuel L. Volchenboum³, Katherine Chou¹, Michael Pearson¹, Srinivasan Madabushi¹, Nigam H. Shah¹, Atul J. Butte², Michael D. Howell¹, Claire Cui¹, Greg S. Corrado¹ and Jeffrey Dean¹

Predictive modeling with electronic health record (EHR) data is anticipated to drive personalized medicine and improve healthcare quality. Constructing predictive statistical models typically requires extraction of curated predictor variables from normalized EHR data, a labor-intensive process that discards the vast majority of information in each patient's record. We propose a representation of patients' entire raw EHR records based on the Fast Healthcare Interoperability Resources (FHIR) format. We demonstrate that deep learning methods using this representation are capable of accurately predicting multiple medical events from multiple centers without site-specific data harmonization. We validated our approach using de-identified EHR data from two US academic medical centers with 216,221 adult patients hospitalized for at least 24 h. In the sequential format we propose, this volume of EHR data unrolled into a total of 46,864,534,945 data points, including clinical notes. Deep learning models achieved high accuracy for

health data

© HBR.ORG

DECISION SUPPORT SYSTEMS



DECISION SUPPORT SYSTEMS

LEVELS OF AUTOMATION OF DECISION AND ACTION SELECTION

HIGH

10. The computer decides everything, acts autonomously, ignoring the human.

9. informs the human only if it, the computer, decides to

8. informs the human only if asked, or

7. executes automatically, then necessarily informs the human, and

6. allows the human a restricted time to veto before automatic execution, or

5. executes that suggestion if the human approves, or

4. suggests one alternative

3. narrows the selection down to a few, or

2. The computer offers a complete set of decision/action alternatives, or

LOW

1. The computer offers no assistance: human must take all decisions and actions.

DECISION SUPPORT SYSTEMS

LEVELS OF AUTOMATION OF DECISION AND ACTION SELECTION

- HIGH
10. The computer decides everything, acts autonomously, ignoring the human.
 9. informs the human only if it, the computer, decides to
 8. informs the human only if asked, or
 7. executes automatically, then necessarily informs the human, and
 6. allows the human a restricted time to veto before automatic execution, or
 5. executes that suggestion if the human approves, or
 4. suggests one alternative
 3. narrows the selection down to a few, or
 2. The computer offers a complete set of decision/action alternatives, or
- LOW
1. The computer offers no assistance: human must take all decisions and actions.

HCI

DECISION SUPPORT SYSTEMS

LEVELS OF AUTOMATION OF DECISION AND ACTION SELECTION

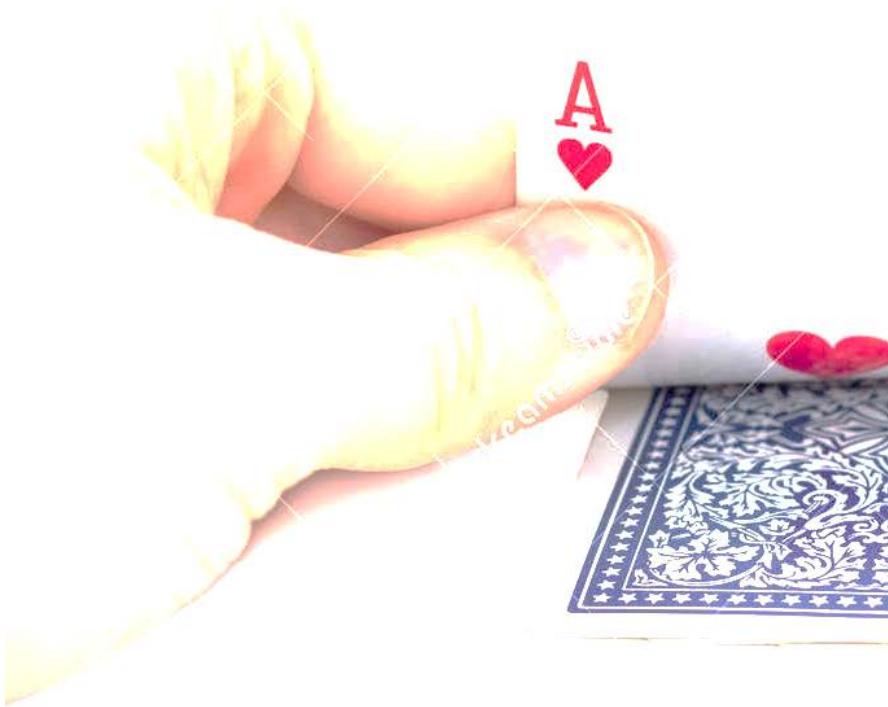
- HIGH
10. The computer decides everything, acts autonomously, ignoring the human.
 9. informs the human only if it, the computer, decides to
 8. informs the human only if asked, or
 7. executes automatically, then necessarily informs the human, and
 6. allows the human a restricted time to veto before automatic execution, or
 5. executes that suggestion if the human approves, or
 4. suggests one alternative
 3. narrows the selection down to a few, or
 2. The computer offers a complete set of decision/action alternatives, or
- LOW
1. The computer offers no assistance: human must take all decisions and actions.



PREDIZIONE DELL'ESITO DI UN INTERVENTO CHIRURGICO



PREDIZIONE DELL'ESITO DI UN INTERVENTO CHIRURGICO



HIP:: MENTAL SCORE :: 3 MONTHS

Prediction of the mental score from the SF12 questionnaire (MCS) of hip surgery (first intervention) after 3 months.

Selected features:

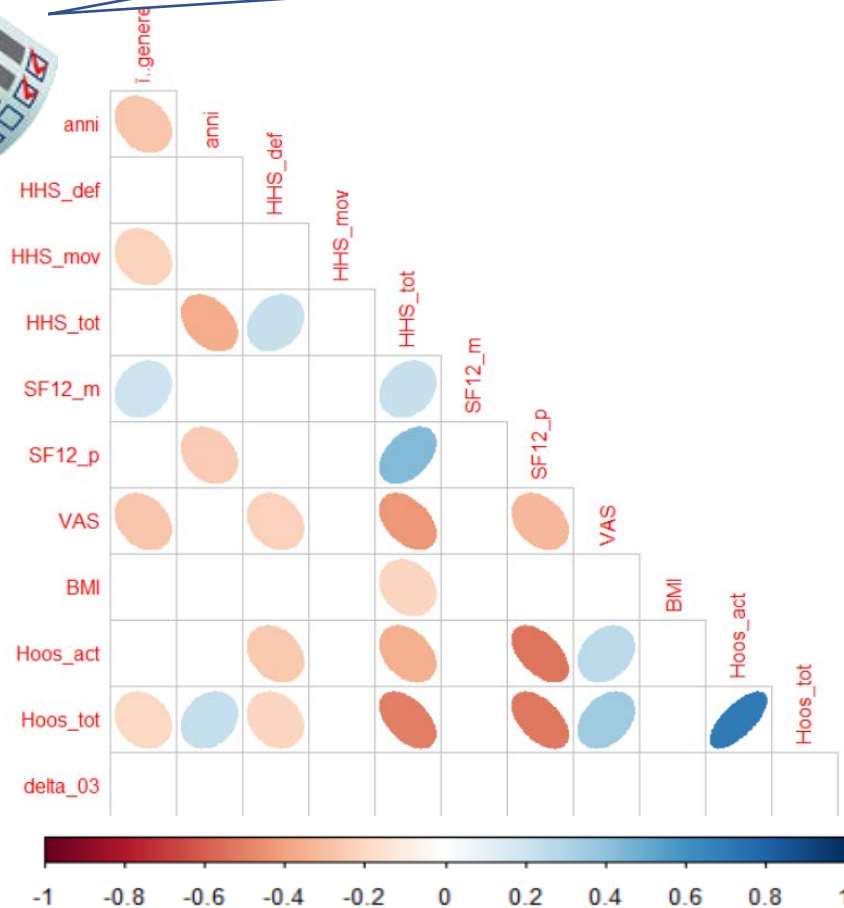
1. Age,
2. gender,
3. pre-op HHS deformity,
4. per-op HSS movement,
5. per-op HSS total
6. pre-op SF12 mental score,
7. pre-op SF12 physical score.
8. Pre-op pain VAS
9. Pre-op BMI
10. Preop Hoos-ps

Target variable:

- Delta SF12 mental score (MCS) 0-3 months

PREDIZIONE DELL'ESITO DI UN INTERVENTO CHIRURGICO

Nessun imbroglio: le variabili di input non hanno alcuna correlazione con l'output.



HIP:: MENTAL SCORE :: 3 MONTHS

Prediction of the mental score from the SF12 questionnaire (MCS) of hip surgery (first intervention) after 3 months.

Selected features:

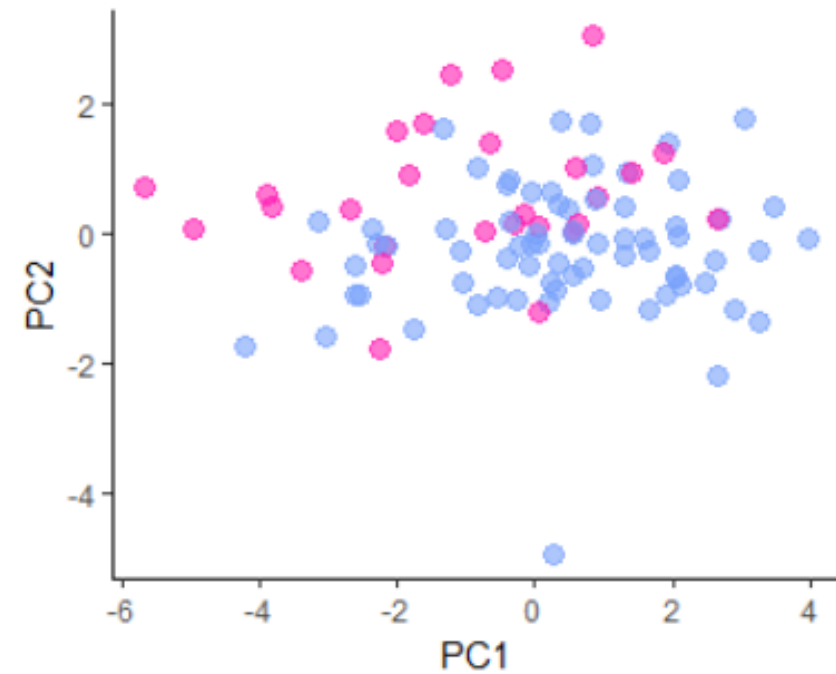
1. Age,
2. gender,
3. pre-op HSS deformity,
4. per-op HSS movement,
5. per-op HSS total
6. pre-op SF12 mental score,
7. pre-op SF12 physical score.
8. Pre-op pain VAS
9. Pre-op BMI
10. Preop Hoos-ps activity

Target variable:

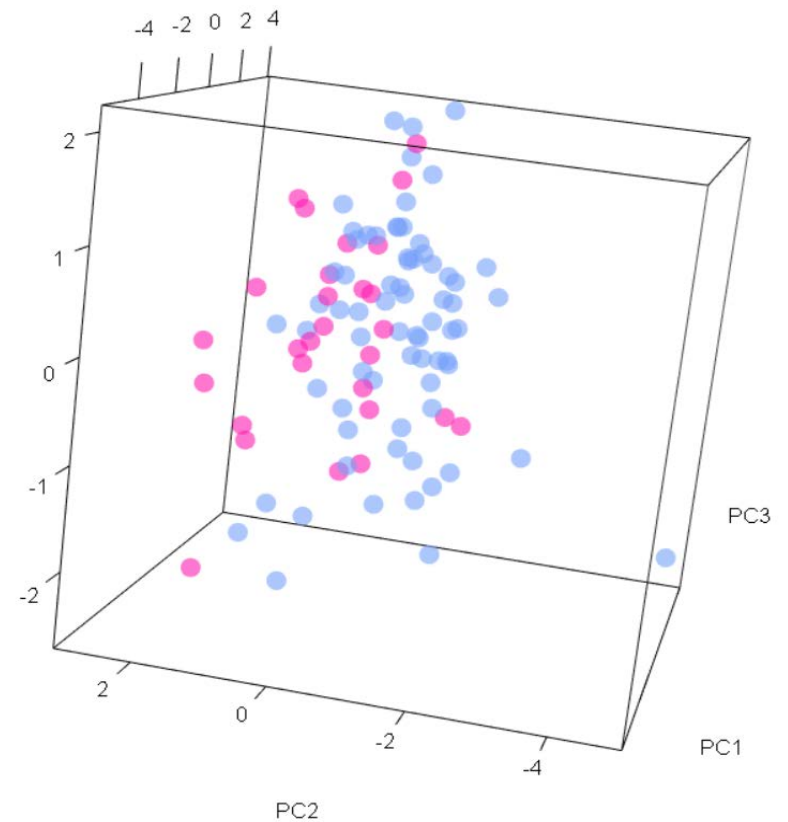
- Delta SF12 mental score (MCS) 0-3 months

PREDIZIONE DELL'ESITO DI UN INTERVENTO CHIRURGICO

Nessun imbroglio: peggiorati e migliorati sono effettivamente "mischianti".



Classi
● $(-\text{Inf}, 0]$
● $(0, \text{Inf}]$



PREDIZIONE DELL'ESITO DI UN INTERVENTO CHIRURGICO

Il modello più accurato per questo problema è risultato il Random Forest.



weka.classifiers.trees.RandomForest

About

Class for constructing a forest of random trees.

bagSizePercent

batchSize

breakTiesRandomly ▼

calcOutOfBag ▼

computeAttributeImportance ▼

debug ▼

doNotCheckCapabilities ▼

maxDepth

numDecimalPlaces

numExecutionSlots

numFeatures

numIterations

outputOutOfBagComplexityStatistics ▼

printClassifiers ▼

seed

storeOutOfBagPredictions ▼

HIP:: MENTAL SCORE :: 3 MONTHS

Prediction of the mental score from the SF12 questionnaire (MCS) of hip surgery (first intervention) after 3 months.

Selected features:

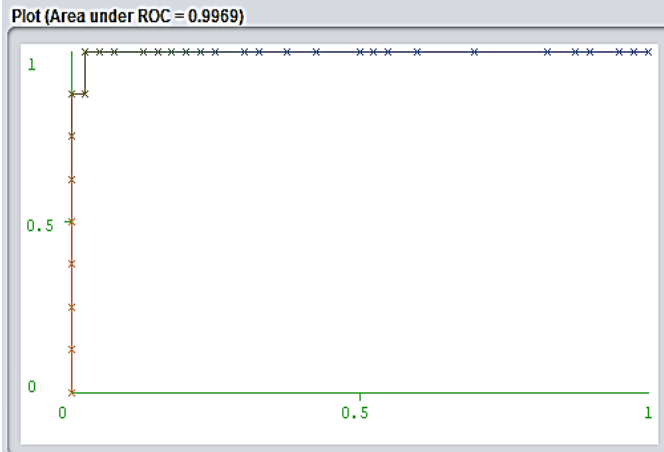
- | | |
|--------------------------|--------------------------------|
| 1. Age, | 6. pre-op SF12 mental score, |
| 2. gender, | 7. pre-op SF12 physical score. |
| 3. pre-op HHS deformity, | 8. Pre-op pain VAS |
| 4. per-op HSS movement, | 9. Pre-op BMI |
| 5. per-op HSS total | 10. Preop Hoos-ps activity |

Target variable:

- Delta SF12 mental score (MCS) 0-3 months

PREDIZIONE DELL'ESITO DI UN INTERVENTO CHIRURGICO

Il modello più accurato per questo problema è risultato il Random Forest.



Accuracy: 97.9%
TP rate: 97.4%
FP rate: 0.04%
Precision: 98.1%
Recall: 97.9%
F-score: 98.0% (on test set)

HIP:: MENTAL SCORE :: 3 MONTHS

Prediction of the mental score from the SF12 questionnaire (MCS) of hip surgery (first intervention) after 3 months.

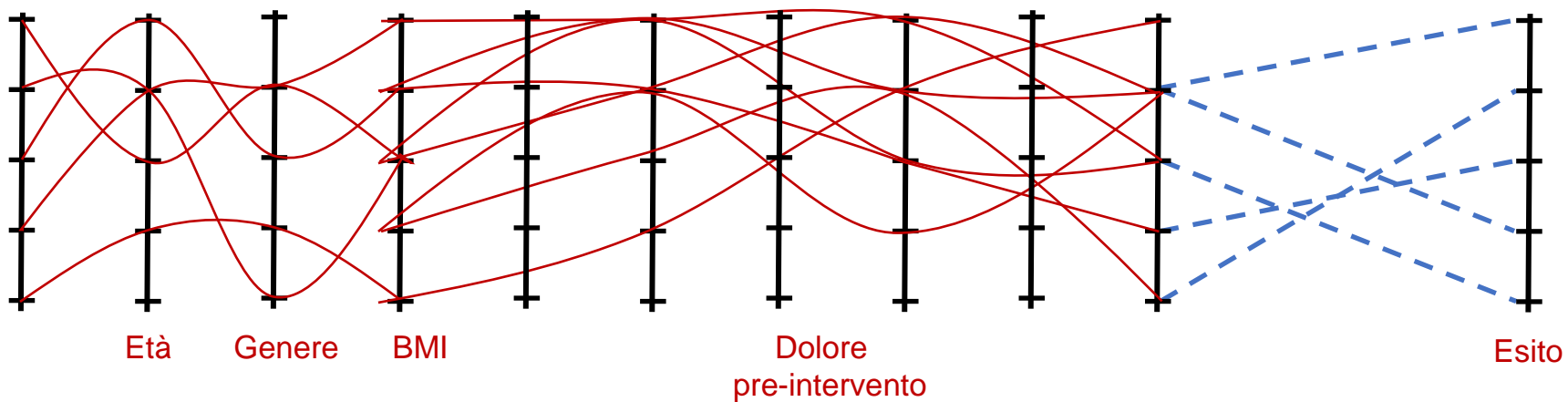
Selected features:

1. Age,	6. pre-op SF12 mental score,
2. gender,	7. pre-op SF12 physical score.
3. pre-op HHS deformity,	8. Pre-op pain VAS
4. pre-op HHS deformity,	9. e-BMI
5. pre-op total	10. pre-op physical activity

99%

Target variable:
➤ Delta SF12 mental score (MCS) 0-3 months

A partire da solo 10 variabili iniziali



99%

DECISION SUPPORT SYSTEMS

LEVELS OF AUTOMATION OF DECISION AND ACTION SELECTION

- HIGH
10. The computer decides everything, acts autonomously, ignoring the human.
 9. informs the human only if it, the computer, decides to
 8. informs the human only if asked, or
 7. executes automatically, then necessarily informs the human, and
 6. allows the human a restricted time to veto before automatic execution, or
 5. executes that suggestion if the human approves, or
 4. suggests one alternative
 3. narrows the selection down to a few, or
 2. The computer offers a complete set of decision/action alternatives, or
 1. The computer offers no assistance: human must take all decisions and actions.

Deskilling?


Augmentation?



DECISION SUPPORT SYSTEMS

C'è un gran dibattito sul potenziale di sostituzione degli esseri umani per via dell'adozione di sistemi di AI (tipicamente basati su Machine Learning).

DECISION SUPPORT SYSTEMS



C'è un gran dibattito sul potenziale di sostituzione degli esseri umani per via dell'adozione di sistemi di AI (tipicamente basati su Machine Learning).

Ma forse è meglio capire che competenze perdiamo se ci affidiamo troppo a questi strumenti.

AUTOMATION BIAS



AUTOMATION BIAS

L'Automation bias si verifica quando un essere umano si affida eccessivamente ad un supporto computazionale e tende a non prendere più decisioni sulla base di una analisi approfondita dei dati disponibili (“si fida”).

Potenziali conseguenze inattese dell'uso di sistemi di intelligenza artificiale oracolari in medicina

FEDERICO CABITZA^{1,2*}, CAMILLA ALDERIGHI^{3*}, RAFFAELE RASOINI^{3,4*}, GIAN FRANCO GENSINI^{3*}

¹IRCCS Istituto Ortopedico Galeazzi, Milano; ² Dipartimento di Informatica, Sistemistica e Comunicazione, Università di Milano-Bicocca, Milano; ³CESMAV - Centro Studi Medicina Avanzata, Firenze; ⁴IFCA Istituto Fiorentino di Cura e Assistenza, Firenze.

*Florence EBM/Renaissance Group.

Pervenuto il 16 agosto 2017. Accettato dopo revisione il 31 agosto 2017.

Riassunto. I sistemi di supporto decisionale basati sul *machine learning* (ML) in medicina stanno raccogliendo un crescente interesse grazie a recenti pubblicazioni che ne hanno evidenziato l'elevata accuratezza diagnostica in specifici contesti clinici. Tuttavia, agli ipotetici vantaggi derivanti dall'applicazione dei sistemi di intelligenza artificiale in campo medico, vanno criticamente affiancati alcuni potenziali inconvenienti. Alla luce dell'attuale mancanza di studi sugli effetti collaterali dell'applicazione di questi nuovi supporti decisionali nella pratica medica,

"Handle with care": about the potential unintended consequences of oracular artificial intelligence systems in medicine.

Summary. Decisional support systems based on *machine learning* (ML) in medicine are gaining a growing interest as some recent articles have highlighted the high diagnostic accuracy exhibited by these systems in specific medical contexts. However, it is implausible that any potential advantage can be obtained without some potential

VIEWPOINT

Unintended Consequences of Machine Learning in Medicine

Federico Cabitza, PhD
Department of
Informatics, University
of Milano-Bicocca,
Milan, Italy; and IRCCS
Istituto Ortopedico
Galeazzi, Milan, Italy.

Raffaele Rasoini, MD
Centro Studi Medicina
Avanzata, Florence,
Italy.

**Gian Franco Gensini,
MD**
Centro Studi Medicina
Avanzata, Florence,
Italy.

Over the past decade, machine learning techniques have made substantial advances in many domains. In health care, global interest in the potential of machine learning has increased; for example, a deep learning algorithm has shown high accuracy in detecting diabetic retinopathy.¹ There have been suggestions that machine learning will drive changes in health care within a few years, specifically in medical disciplines that require more accurate prognostic models (eg, oncology) and those based on pattern recognition (eg, radiology and pathology).

However, comparative studies on the effectiveness of machine learning-based decision support systems (ML-DSS) in medicine are lacking, especially regarding the effects on health outcomes. Moreover, the introduction of new technologies in health care has not always been straightforward or without unintended and

the expense of other elements that are more difficult or impossible to easily describe. Relying on ML-DSS requires considering digital data as reliable and complete representations of the phenomena that these data are supposed to render in a discrete and trustworthy form. This may be a problem when the clinical context is not represented, particularly if physicians lose awareness of the existence of clinical elements that are not included in the clinical record.

Such lack of information may lead to partial or misleading interpretations of ML-DSS diagnostics and therapeutic or prognostic outputs. It also could lead to reduced interest in and decreased ability to perform holistic evaluations of patients, with loss of valuable and irreducible aspects of the human experience such as psychological, relational, social, and organizational issues.

#13

of 382 outputs



Altmetric has tracked 11,422,063 research outputs across all sources so far. Compared to these this one has done particularly well and is in the 99th percentile: it's **in the top 5% of all research outputs ever tracked** by Altmetric.

Recenti Prog Med 2017; 108: 397-401

VIEWPOINT

Unintended Consequences of Machine Learning in Medicine

Federico Cabitza, PhD
Department of Informatics, University of Milano-Bicocca, Milan, Italy; and IRCCS Istituto Ortopedico Galeazzi, Milan, Italy.

Raffaele Rasoini, MD
Centro Studi Medicina Avanzata, Florence, Italy.

Gian Franco Gensini, MD
Centro Studi Medicina Avanzata, Florence, Italy.

Over the past decade, machine learning techniques have made substantial advances in many domains. In health care, global interest in the potential of machine learning has increased; for example, a deep learning algorithm has shown high accuracy in detecting diabetic retinopathy.¹ There have been suggestions that machine learning will drive changes in health care within a few years, specifically in medical disciplines that require more accurate prognostic models (eg, oncology) and those based on pattern recognition (eg, radiology and pathology).

However, comparative studies on the effectiveness of machine learning-based decision support systems (ML-DSS) in medicine are lacking, especially regarding the effects on health outcomes. Moreover, the introduction of new technologies in health care has not always been straightforward or without unintended and

the expense of other elements that are more difficult or impossible to easily describe. Relying on ML-DSS requires considering digital data as reliable and complete representations of the phenomena that these data are supposed to render in a discrete and trustworthy form. This may be a problem when the clinical context is not represented, particularly if physicians lose awareness of the existence of clinical elements that are not included in the clinical record.

Such lack of information may lead to partial or misleading interpretations of ML-DSS diagnostics and therapeutic or prognostic outputs. It also could lead to reduced interest in and decreased ability to perform holistic evaluations of patients, with loss of valuable and irreducible aspects of the human experience such as psychological, relational, social, and organizational issues.

Dovrebbe poi divenire ben radicata, tra gli informatici e tra gli esperti di ML, l'esigenza di non trascurare l'incertezza insita nell'interpretazione di ogni fenomeno in medicina e anzi prioritario considerare tale incertezza, e valorizzarla, nell'ambito degli algoritmi di definizione dei modelli predittivi e nella interpretazione dei loro risultati¹⁷.

Andrebbe inoltre evidenziato e studiato il pericolo di sovra-affidamento e di eccessiva dipendenza da sistemi di accuratezza che definiamo "oracolare", cioè molto elevata ma non associata a spiegazioni esplicite e significative per gli operatori coinvolti⁸, con gli effetti secondari di deskilling e desensibilizzazione al contesto clinico.

AUTOMATION BIAS

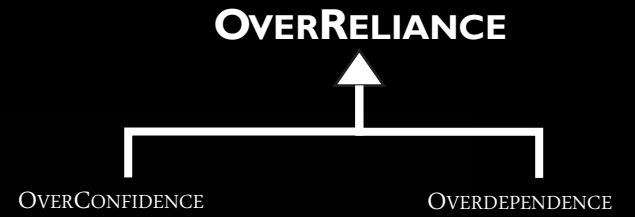
“Automation bias è stato osservato sia nell’inesperto che nell’esperto, non può essere evitato con una adeguata formazione o insieme di istruzioni, e può condizionare la decisione, inducendo errori sia di omissione che di commissione (overdiagnosis), da parte di individui come di team.” *

AUTOMATION BIAS

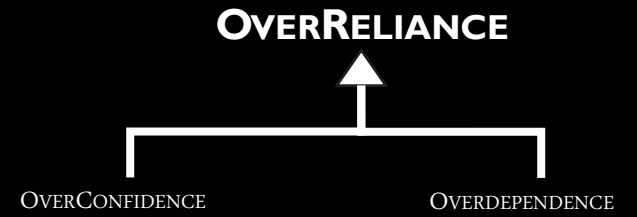
OVERRELIANCE



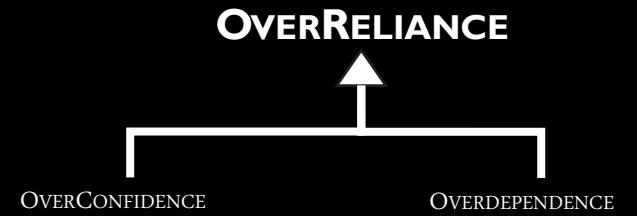
AUTOMATION BIAS



AUTOMATION BIAS



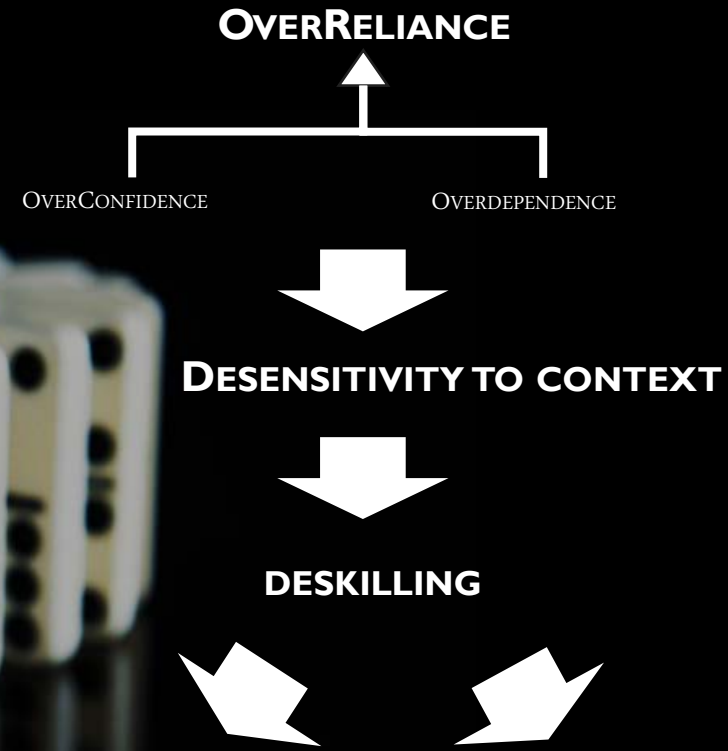
AUTOMATION BIAS



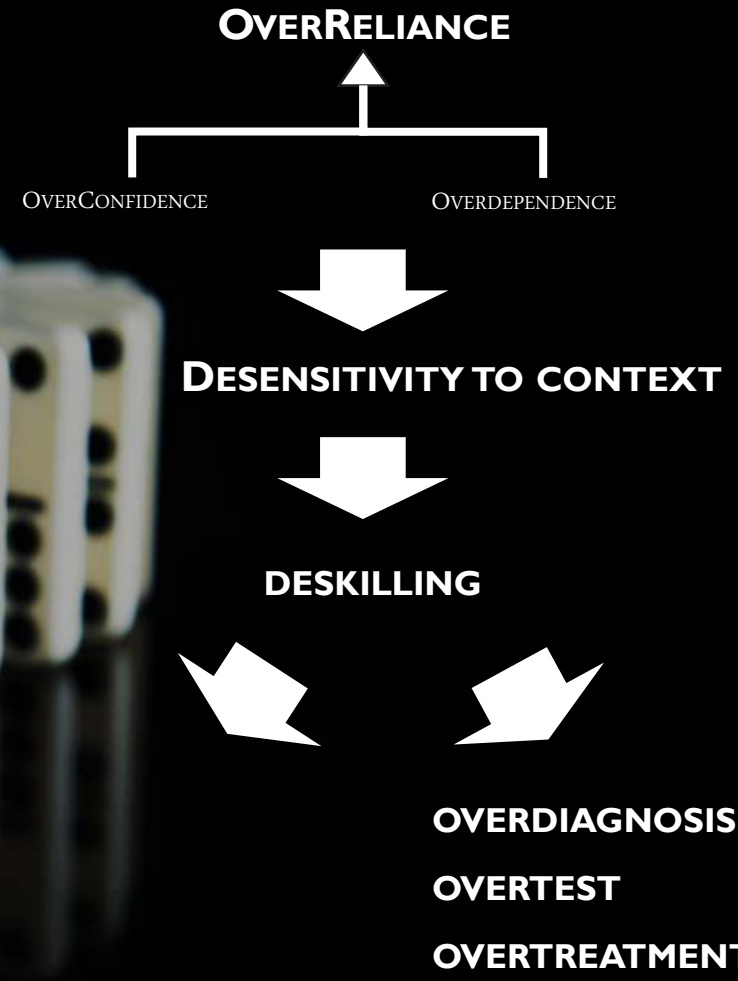
↓
DESENSITIVITY TO CONTEXT

↓
DESKILLING

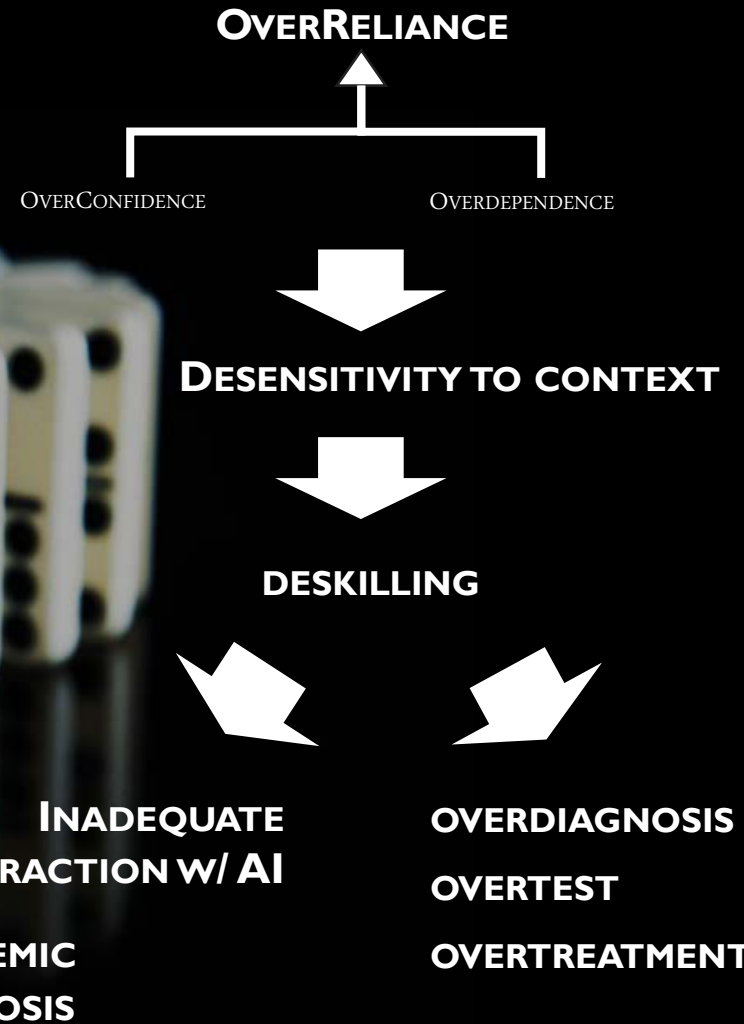
AUTOMATION BIAS



AUTOMATION BIAS



AUTOMATION BIAS



Gli algoritmi AI sono come farmaci?

ALGORITHMS ARE THE NEW DRUGS? REFLECTIONS FOR A CULTURE OF IMPACT ASSESSMENT AND VIGILANCE

Hugh Harvey

*Kheiron Medical Technologies
40 islington High street, London, N1 8EQ, UK*

Federico Cabitza

*Università degli Studi di Milano-Bicocca
IRCCS Istituto Ortopedico Galeazzi
Viale Sarca 336, 20126 Milano, Italy*



ABSTRACT

In this paper we propose that Artificially Intelligent (AI) systems in medicine are *the new drugs*. Through study and comparison of the existing corporate pharmaceutical models, we advocate a new approach in how the performance of AI solutions (both for diagnostic and prognostic aims) is measured and promoted, on how the impact of these solutions is assessed and evaluated in terms of both outcome and safety (about which the new term *technovigilance* has been introduced), and how education and a better awareness of the impact of these systems on health care by multiple stakeholders should be promoted. Any drug must be evaluated in terms of efficacy and safety, but also in terms of real-world effectiveness, efficiency and cost-effectiveness, that is in terms of its role as a component of a wider and more complex socio-technical system where it must fit a broader



MCCSIS 2018

MADRID, SPAIN

17 - 20 JULY

MULTI CONFERENCE ON COMPUTER SCIENCE AND INFORMATION SYSTEMS

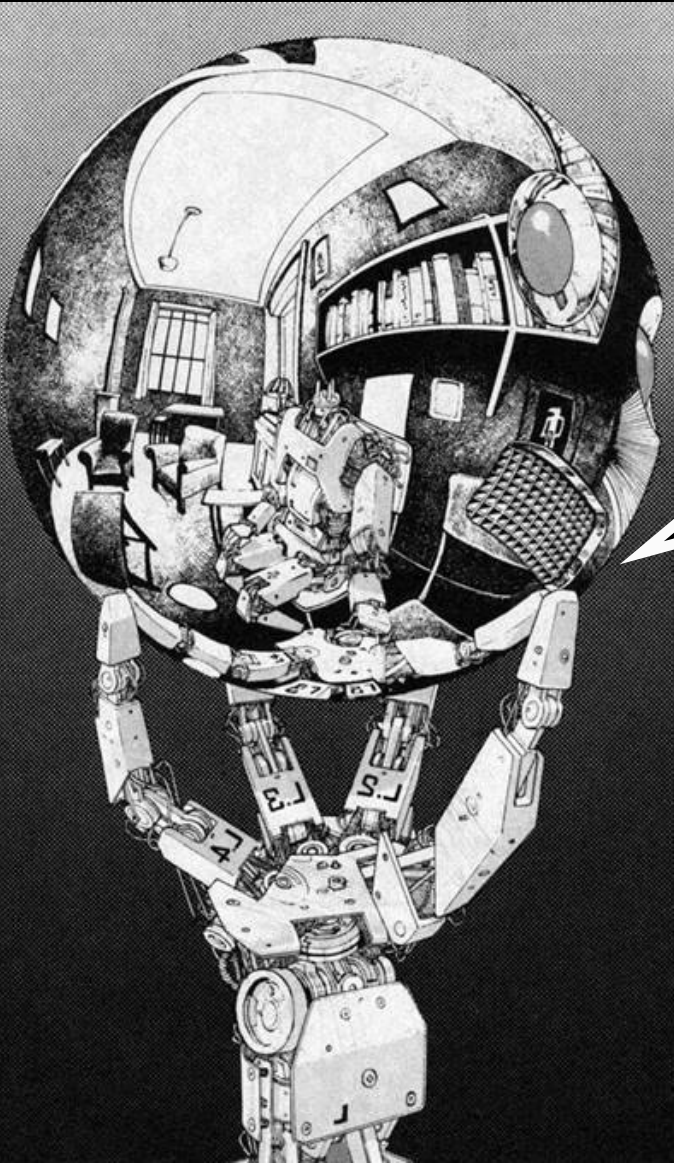


Gli algoritmi AI sono come farmaci?

Sì!

ALGORITHMS ARE THE NEW DRUGS?
REFLECTIONS FOR A CULTURE OF IMPACT
ASSESSMENT AND VIGILANCE

- ✓ Sì, perché la loro adozione può apportare un potenziale di benefici.
- ✓ Sì, perché questo potenziale **DEVE** essere valutato in condizioni real-world (confrontando end point primari e secondari in bracci sperimentali omogenei).
- ✓ Sì, perché la loro promozione e affermazioni su questo potenziale **DEVONO** essere regolamentati.
- ✓ Sì perché bisogna essere consapevoli degli effetti collaterali (tra cui l'automation bias e il deskilling).



GRAZIE!



cabitza @ disco.unimib.it



@cabitza

These slides are not intended nor produced to be made publicly available. However, I share them to a limited readership (those who possess the Web address where the slides are stored) in an effort of disseminating scientific ideas and reflections, with no intended aim of economic gain, and to get feedback and comments on those ideas, which I express under a CC-BY 4.0 license.

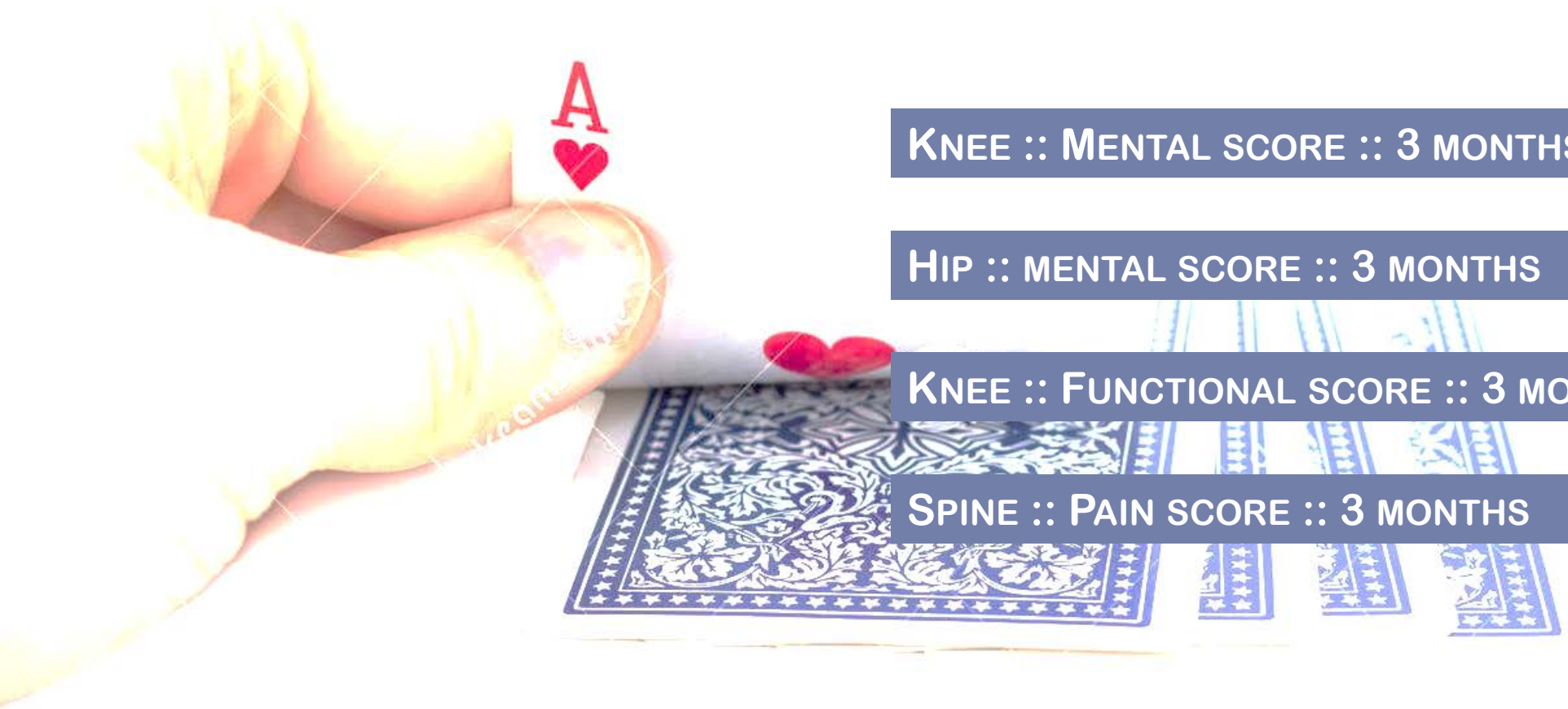
I took the greatest possible care to identify all image copyright holders correctly. However, if I have omitted to do so in some individual instances, I would be most grateful if these copyright holders would inform me forthwith. It is my policy to immediately remove, upon notification and identification, any specific image displayed on this website for which the copyright holder deems the fair use (see below) cannot be associated. Upon request, I will remove immediately the content and update the content accordingly.

FAIR USE NOTICE

These slides may contain copyrighted (©) material the use of which has not always been specifically authorized by the copyright owner. Such material is made available to advance understanding of ecological, political, historical, human rights, economic, artistic expression, democracy, scientific, moral, ethical, and social justice issues, etc. It is believed that this constitutes a 'fair use' of any such copyrighted material as provided for in section 107 of the US Copyright Law. In accordance with Title 17 U.S.C. Section 107, this material is distributed without profit to those who have expressed a prior general interest in receiving similar information for research and educational purposes.

** COPYRIGHT NOTICE** In accordance with Title 17 U.S.C. Se

OUTCOME PREDICTION



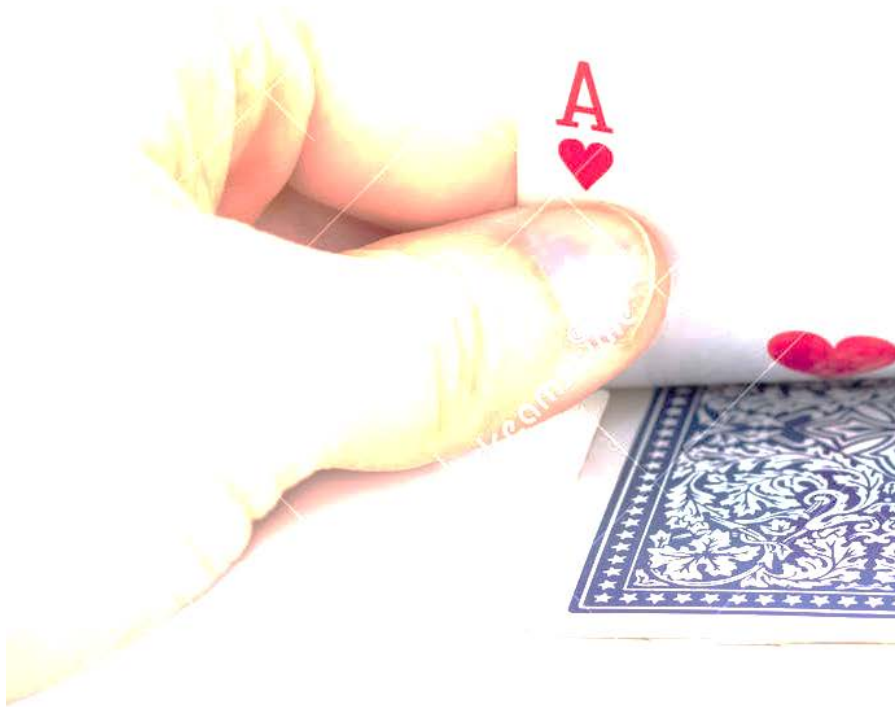
KNEE :: MENTAL SCORE :: 3 MONTHS

HIP :: MENTAL SCORE :: 3 MONTHS

KNEE :: FUNCTIONAL SCORE :: 3 MONTHS

SPINE :: PAIN SCORE :: 3 MONTHS

OUTCOME PREDICTION



HIP:: MENTAL SCORE :: 3 MONTHS

Prediction of the mental score from the SF12 questionnaire (MCS) of hip surgery (first intervention) after 3 months.

Selected features:

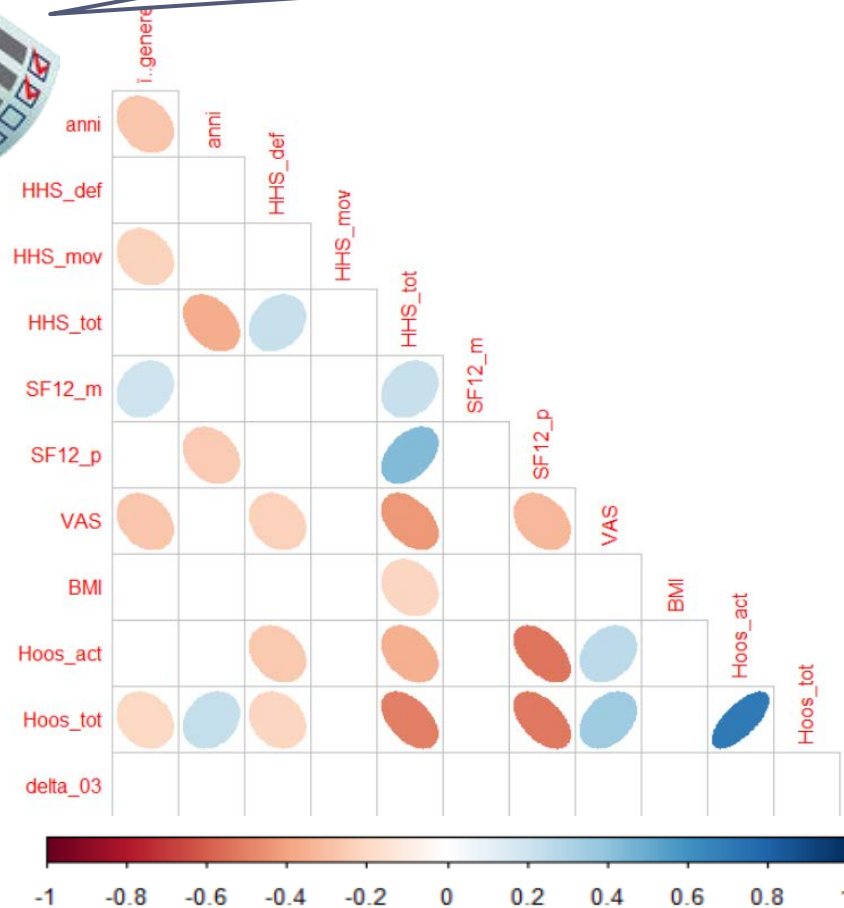
- | | |
|--------------------------|--------------------------------|
| 1. Age, | 6. pre-op SF12 mental score, |
| 2. gender, | 7. pre-op SF12 physical score. |
| 3. pre-op HHS deformity, | 8. Pre-op pain VAS |
| 4. per-op HSS movement, | 9. Pre-op BMI |
| 5. per-op HSS total | 10. Preop Hoos-ps activity |

Target variable:

- Delta SF12 mental score (MCS) 0-3 months

OUTCOME PREDICTION

No cheating: input variables do not correlate with the target variable.



HIP:: MENTAL SCORE :: 3 MONTHS

Prediction of the mental score from the SF12 questionnaire (MCS) of hip surgery (first intervention) after 3 months.

Selected features:

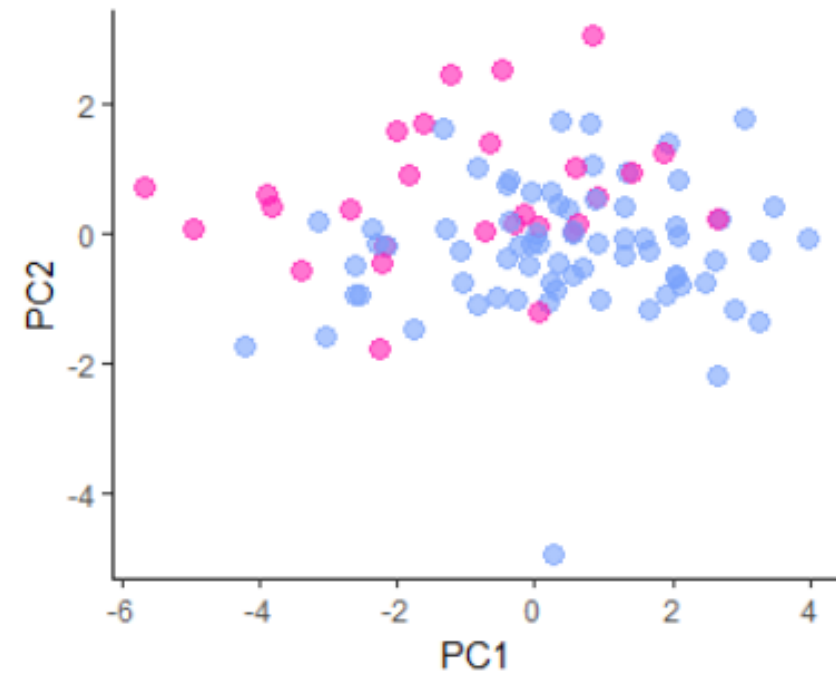
- | | |
|--------------------------|--------------------------------|
| 1. Age, | 6. pre-op SF12 mental score, |
| 2. gender, | 7. pre-op SF12 physical score. |
| 3. pre-op HSS deformity, | 8. Pre-op pain VAS |
| 4. per-op HSS movement, | 9. Pre-op BMI |
| 5. per-op HSS total | 10. Preop Hoos-ps activity |

Target variable:

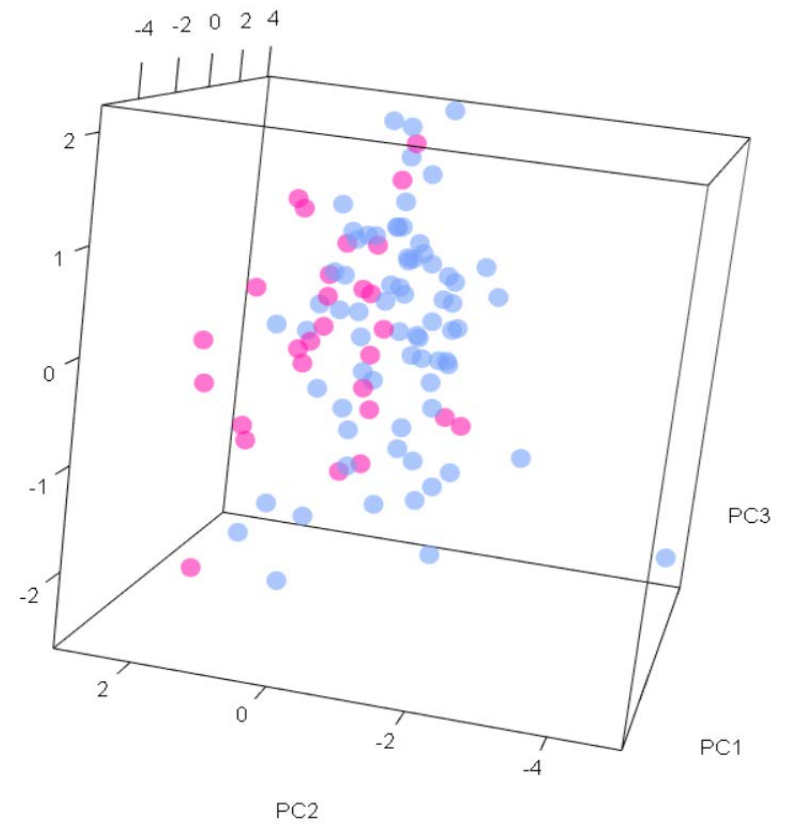
- Delta SF12 mental score (MCS) 0-3 months

OUTCOME PREDICTION

No cheating: worsened patients and those who got better look “mixed”.



Classi
● $(-\text{Inf}, 0]$
● $(0, \text{Inf}]$



OUTCOME PREDICTION

The most accurate model is the Random Forest.



weka.classifiers.trees.RandomForest

About

Class for constructing a forest of random trees.

bagSizePercent

batchSize

breakTiesRandomly ▼

calcOutOfBag ▼

computeAttributeImportance ▼

debug ▼

doNotCheckCapabilities ▼

maxDepth

numDecimalPlaces

numExecutionSlots

numFeatures

numIterations

outputOutOfBagComplexityStatistics ▼

printClassifiers ▼

seed

storeOutOfBagPredictions ▼

HIP:: MENTAL SCORE :: 3 MONTHS

Prediction of the mental score from the SF12 questionnaire (MCS) of hip surgery (first intervention) after 3 months.

Selected features:

- | | |
|--------------------------|--------------------------------|
| 1. Age, | 6. pre-op SF12 mental score, |
| 2. gender, | 7. pre-op SF12 physical score. |
| 3. pre-op HHS deformity, | 8. Pre-op pain VAS |
| 4. per-op HSS movement, | 9. Pre-op BMI |
| 5. per-op HSS total | 10. Preop Hoos-ps activity |

Target variable:

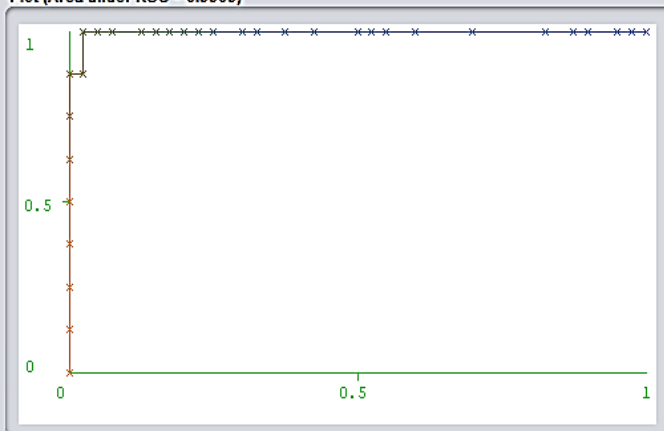
- Delta SF12 mental score (MCS) 0-3 months

OUTCOME PREDICTION

The most accurate discriminative model is the Random Forest. AUROC: 99%



Plot (Area under ROC = 0.9969)



Accuracy: 97.9%
TP rate: 97.4%
FP rate: 0.04%
Precision: 98.1%
Recall: 97.9%
F-score: 98.0% (on test set)

HIP:: MENTAL SCORE :: 3 MONTHS

Prediction of the mental score from the SF12 questionnaire (MCS) of hip surgery (first intervention) after 3 months.

Selected features:

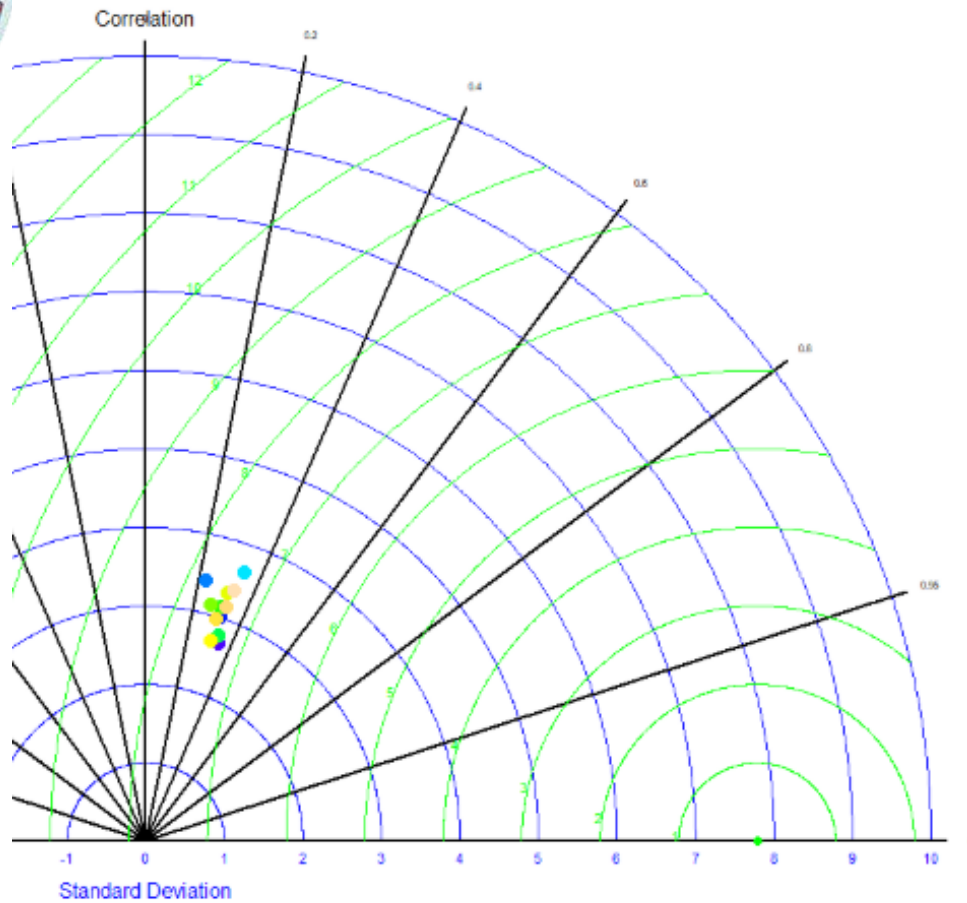
1. Age,	6. pre-op SF12 mental score,
2. gender,	7. pre-op SF12 physical score.
3. pre-op HHS deformity,	8. Pre-op pain VAS
4. pre-op HHS deformity,	9. pre-op BMI
5. pre-op total	10. pre-op physical activity

99%

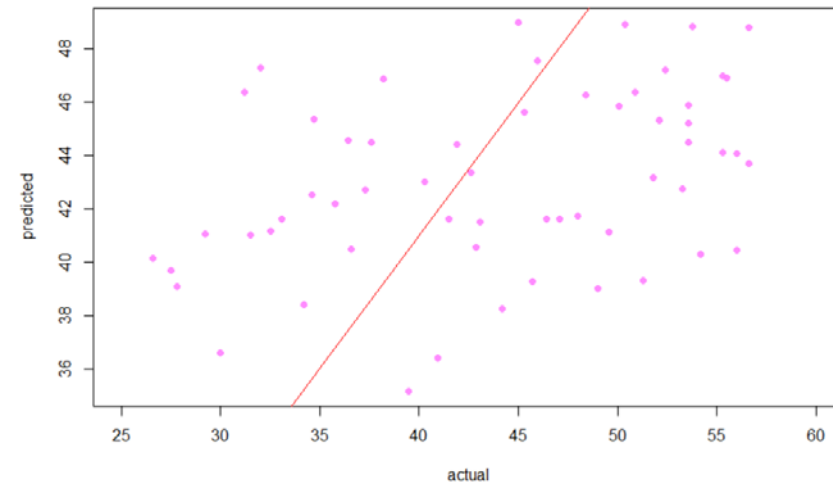
Target variable:
➤ Delta SF12 mental score (MCS) 0-3 months

OUTCOME PREDICTION

The most accurate regression model gives $RMSE = 7.5$.



HIP:: MENTAL SCORE :: 3 MONTHS



7.5 RMSE