



Implementazione di un modello predittivo di risposta alla chemioterapia neoadiuvante

CARLO CAVEDON

UOC FISICA SANITARIA
DIPARTIMENTO DI PATOLOGIA E
DIAGNOSTICA
AZIENDA OSPEDALIERA
UNIVERSITARIA INTEGRATA -
VERONA



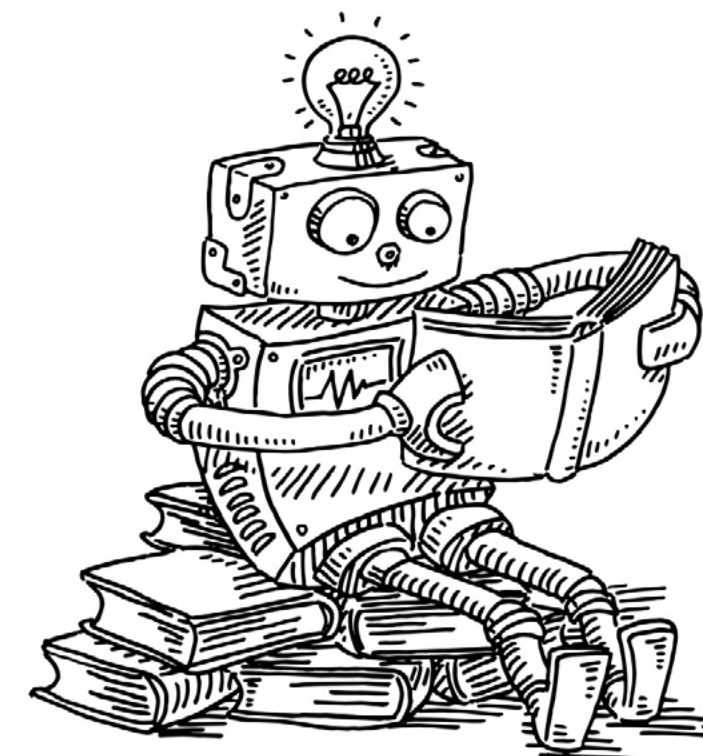
#sanita2030



www.sanita2030.it

Obiettivi

- modello predittivo di **risposta alla chemioterapia neoadiuvante** nel tumore della mammella
- **aumento della fruibilità** di un algoritmo di intelligenza artificiale **nella pratica clinica**
- metodi di **radiomica e AI** per estrarre il massimo delle informazioni dalle immagini radiologiche
- tendenza a breve termine dell'**evoluzione in atto**
- gestione di **sicurezza e qualità** nell'impiego di AI in ambito clinico



Il modello predittivo di risposta completa



3T DCE-MRI Radiomics Improves Predictive Models of Complete Response to Neoadjuvant Chemotherapy in Breast Cancer

Stefania Montemezzi¹, Giulio Benetti², Maria Vittoria Bisighin¹, Lucia Camera¹, Chiara Zerbato¹, Francesca Caumo³, Elena Florio⁴, Sara Zanelli⁴, Michele Zuffante⁵ and Carlo Cavedon^{2*}

¹ Radiology Unit, Azienda Ospedaliera Universitaria Integrata, Verona, Italy, ² Medical Physics Unit, Azienda Ospedaliera Universitaria Integrata, Verona, Italy, ³ Radiology Unit, Istituto Oncologico Veneto – IRCCS, Padova, Italy, ⁴ Pathology Unit, Azienda Ospedaliera Universitaria Integrata, Verona, Italy, ⁵ Nuclear Medicine Unit, Azienda Ospedaliera Universitaria Integrata, Verona, Italy

ORIGINAL RESEARCH
published: 20 April 2021
doi: 10.3389/fonc.2021.630780

- 20 complete responders (pCR) / 40 non-CR
- **radiomica + altri parametri quantitativi**
 - enhancement – type of curve (I-II-III)
 - ADC
 - PET-CT SUV_{max}
 - Ki-67 / ER / PgR / HER-2
 - type of lesion
 - grade / shape / int enhancement
- 5 radiomic features (geometric – 1st order – higher order)
- **classificazione con machine learning**
 - *logistic regression*
 - *support vector machines regression*
 - *random forest*

#sanita2030



www.sanita2030.it



Il modello predittivo di risposta completa

- *feature* di radiomica da sole danno prestazioni inferiori ai parametri istopatologici da soli
- l'inclusione nei modelli di *feature* di radiomica migliora la prestazione di tutti i modelli
- AUC > 0.9 con tutti i classificatori studiati
- => **l'analisi radiomica in MRI fornisce informazioni complementari (non ridondanti)**

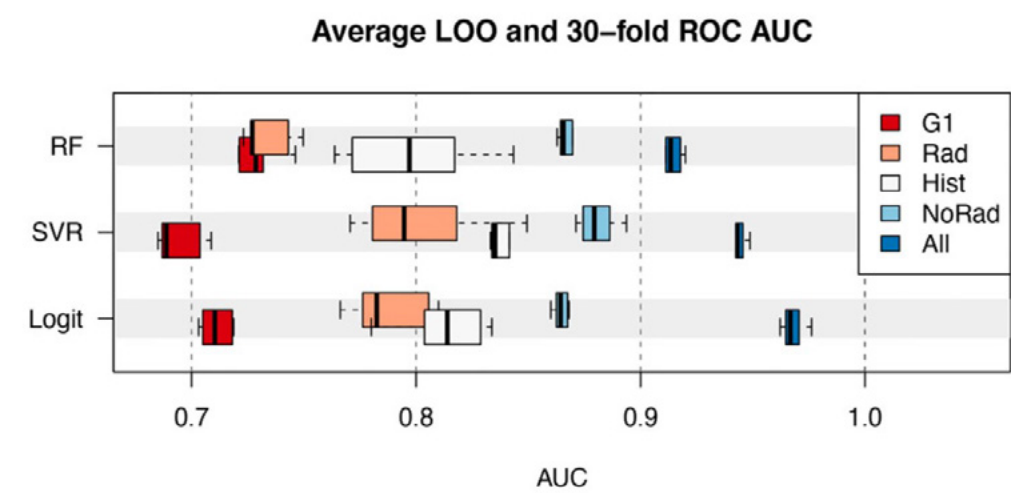


FIGURE 3 | Average AUC of the best 6 models as a function of the class of variables (G1, Rad, Hist, NoRad and All) and classifier (RF, random forest; SVR, Support Vector machines Regression; Logit, Logistic regression). Boxplots represent the median value, interquartile range and extremes.

Traslazione del modello alla clinica

- **obiettivo generale:** facilitare l'introduzione di modelli predittivi basati su metodi di intelligenza artificiale nel flusso di lavoro in ospedale
- **software** che implementa il modello di prestazioni e robustezza migliori
- **obiettivo specifico:** mettere a disposizione della *Breast Unit* uno strumento per la **validazione prospettica** del modello

Traslazione alla clinica: implementazione in un modello distribuito

```
# import tensorflow and keras (tf.keras not "vanilla" Keras)
import tensorflow as tf
from tensorflow import keras

# get data
(train_images, train_labels), (test_images, test_labels) = \
keras.datasets.mnist.load_data()

# setup model
model = keras.Sequential([
    keras.layers.Flatten(input_shape=(28, 28)),
    keras.layers.Dense(128, activation=tf.nn.relu),
    keras.layers.Dense(10, activation=tf.nn.softmax)
])
model.compile(optimizer=tf.train.AdamOptimizer(),
              loss='sparse_categorical_crossentropy',
              metrics=['accuracy'])

# train model
model.fit(train_images, train_labels, epochs=5)

# evaluate
test_loss, test_acc = model.evaluate(test_images, test_labels)
print('test accuracy:', test_acc)

# make predictions
predictions = model.predict(test_images)
```

- acquisizione dei **dati**
- setup del **modello**
 - estrazione *feature*
 - scelta classificatore/regress.
 - scelta ottimizzatore
- **training** del modello
- **test** del modello
- uso del modello per **nuove predizioni**

Implementazione in un modello distribuito

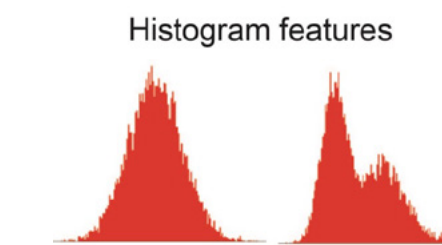


BIRRa Is Real Radiomics



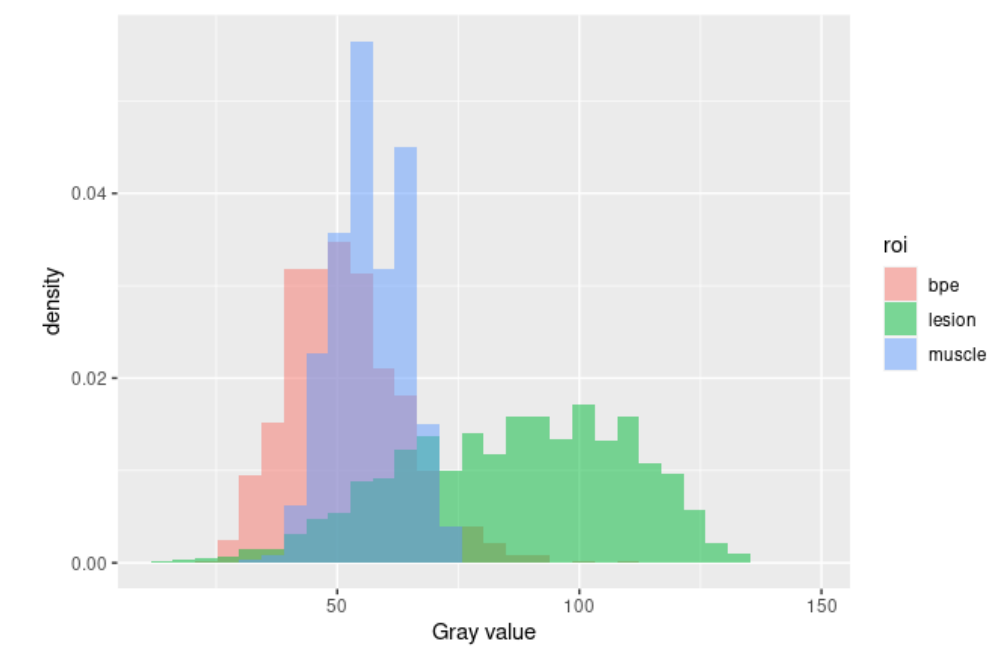
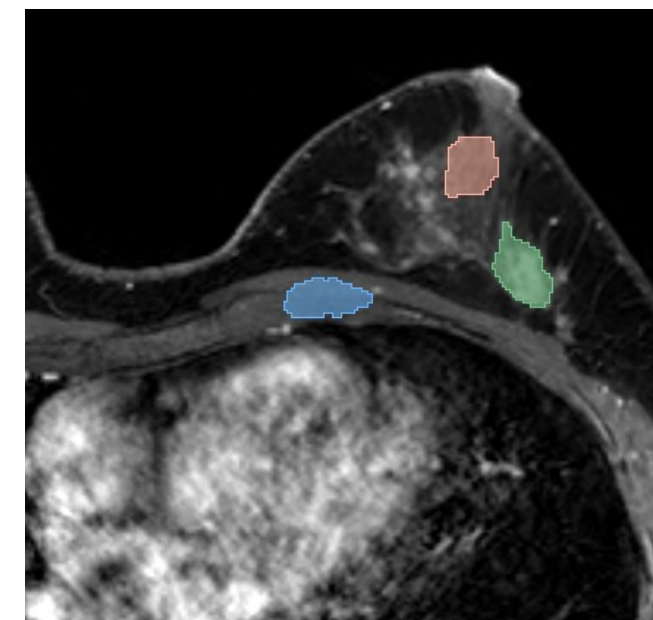
Radiomica: immagini come numeri

- **Features (o metriche) del primo ordine**
 - *intensity-based statistics*
 - caratteristiche dell'istogramma d'immagine
 - ad es. *skewness, kurtosis*
- **Features (o metriche) di forma**
 - ad es. *sphericity*
- **Features (o metriche) di ordine superiore**
 - *texture-based features*
 - raggruppate in *classi* (GLCM, GLRLM...)
 - ad es. *GLRLM_SRE=Short Run Emphasis*



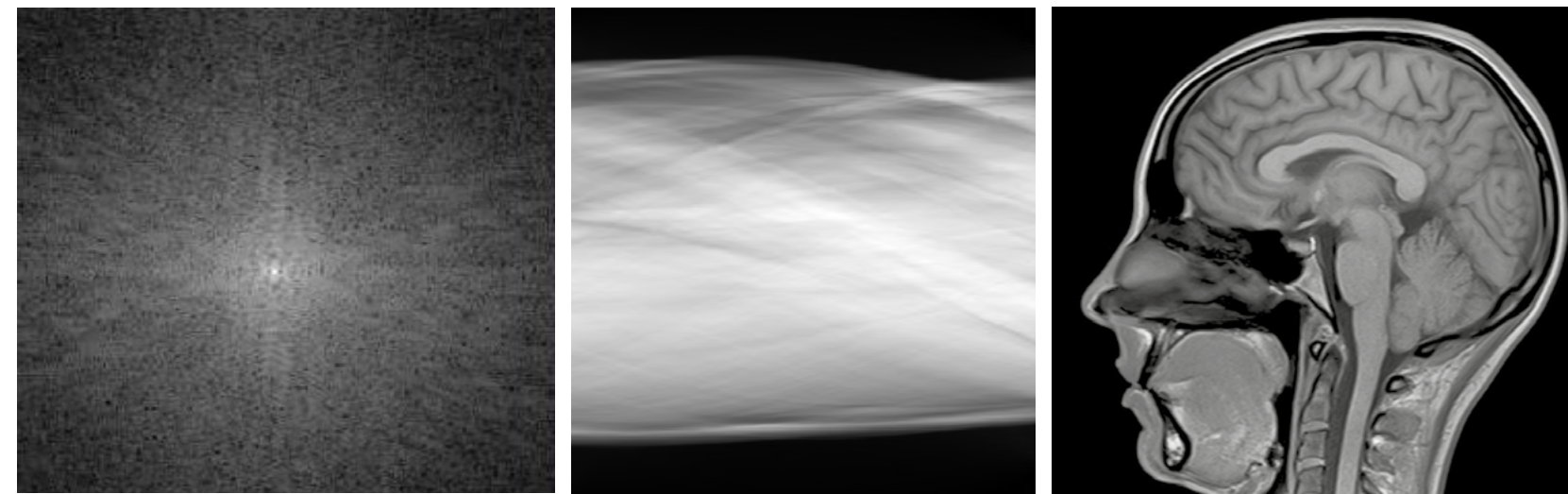
Radiomica: immagini come numeri

- **Features (o metriche) del primo ordine:**
 - *caratteristiche dell'istogramma dei valori di grigio (più correttamente: del contenuto dei pixel/voxel)*



Radiomica: a monte dell'immagine

evoluzione in atto



- **raw data** impiegati da **algoritmi di deep learning**: potenzialità di sfruttare tutta l'informazione presente
- criticità: ridotta possibilità di interpretazione / **validazione umana**

In the Era of Deep Learning, Why Reconstruct an Image at All?

Caroline Chung, MD, MSc, CIP, Jayashree Kalpathy-Cramer, PhD, Michael V. Knopp, MD, David A. Jaffray, PhD



Journal of the American College of Radiology
Volume 18, Issue 1, Part B, January 2021, Pages 170-173



Phys. Med. Biol. 69 (2024) 075015

<https://doi.org/10.1088/1361-6560/ad1e7c>

Physics in Medicine & Biology

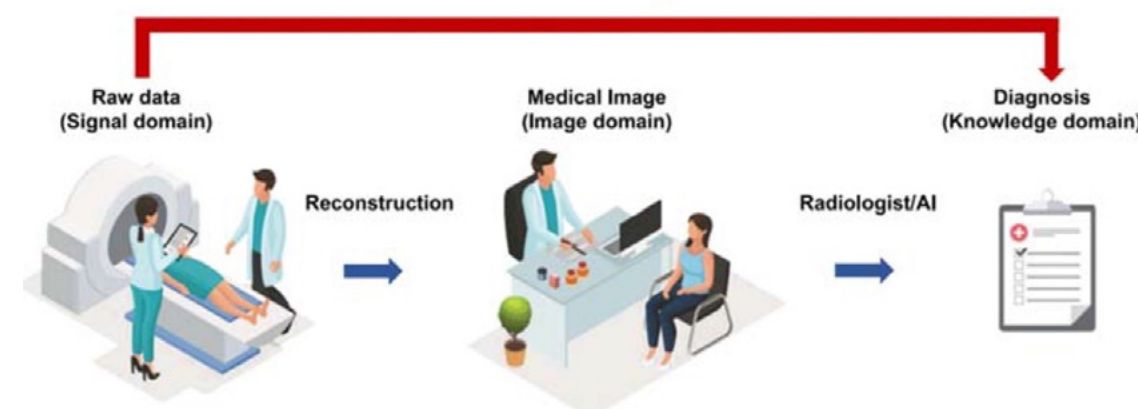


PAPER

Breaking boundaries in radiology: redefining AI diagnostics via raw data ahead of reconstruction

Bingxi He^{1,2,3,9}, Caixia Sun^{1,2,3,9}, Hailin Li^{1,2,3,9}, Yongbo Wang^{4,9}, Yunlang She⁵, Mengmeng Zhao⁵, Mengjie Fang^{1,2,3}, Yongbei Zhu^{1,2,3}, Kun Wang³, Zhenyu Liu³, Ziqi Wei³, Wei Mu^{1,2,3}, Shuo Wang^{1,2,3}, Zhenchao Tang^{1,2,3}, Jingwei Wei³, Lizhi Shao³, Lixia Tong⁶, Feng Huang⁶, Mingze Tang⁷, Yu Guo⁸, Huimao Zhang⁸, Di Dong^{3,4}, Chang Chen^{3,4}, Jianhua Ma^{4,8} and Jie Tian^{1,2,3,9}

Raw data Analysis




#sanita2030



www.sanita2030.it





www.nature.com/scientificreports

SCIENTIFIC REPORTS
nature research

L'INTELLIGENZA ARTIFICIALE IN SANITÀ
TRA ETICA E APPLICAZIONE

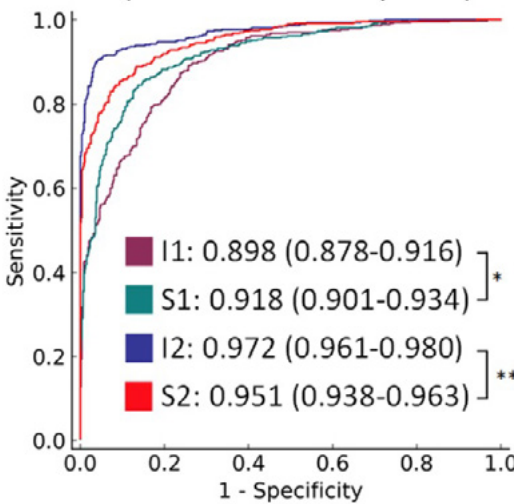
OPEN Machine Friendly Machine Learning: Interpretation of Computed Tomography Without Image Reconstruction

Hyunkwang Lee^{1,2}, Chao Huang¹, Sehyo Yune¹, Shahein H. Tajmir¹, Myeongchan Kim¹ & Synho Do^{1*}

SCIENTIFIC REPORTS | (2019) 9:15540 | <https://doi.org/10.1038/s41598-019-51779-5>

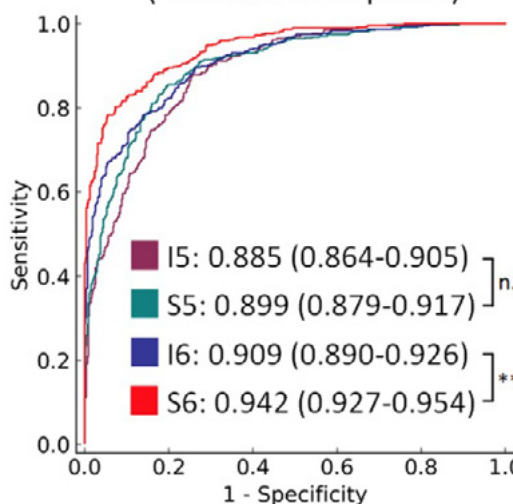
- training su immagini ricostruite e su sinogrammi CT
 - *riconoscimento dettagli anatomici*
 - *identificazione di anomalie*
- prestazioni superiori su immagine se *fully sampled*
- prestazioni superiori su sinogrammi se *sparsely sampled*

Fully sampled
(360 views x 729 pixels)




I1: 0.898 (0.878-0.916)
S1: 0.918 (0.901-0.934)
I2: 0.972 (0.961-0.980)
S2: 0.951 (0.938-0.963)

Sparsely sampled
(40 views x 80 pixels)



I5: 0.885 (0.864-0.905)
S5: 0.899 (0.879-0.917)
I6: 0.909 (0.890-0.926)
S6: 0.942 (0.927-0.954)

#sanita2030



www.sanita2030.it

www.nature.com/scientificreports

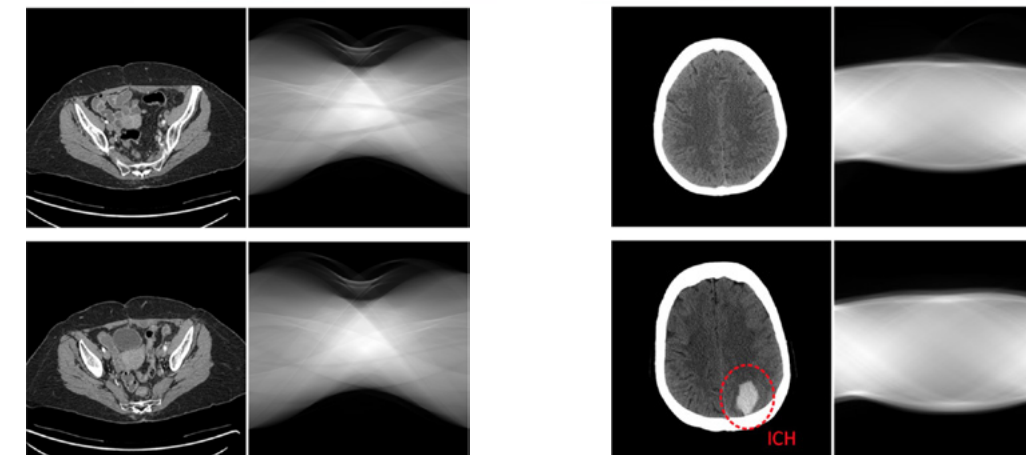
**SCIENTIFIC
REPORTS**
nature research

OPEN Machine Friendly Machine Learning: Interpretation of Computed Tomography Without Image Reconstruction

Hyunkwang Lee^{1,2}, Chao Huang¹, Sehyo Yune¹, Shahein H. Tajmir¹, Myeongchan Kim¹ & Synho Do^{1*}

SCIENTIFIC REPORTS | (2019) 9:15540 | <https://doi.org/10.1038/s41598-019-51779-5>

than conventional networks operating in image-space. As a result, sinogram-space algorithms could be used in field settings for triage (presence of ICH), especially where low radiation dose is desired. These findings also demonstrate another strength of deep learning where it can analyze and interpret sinograms that are virtually impossible for human experts.



#sanita2030



www.sanita2030.it



Le *features* come biomarker nei modelli predittivi

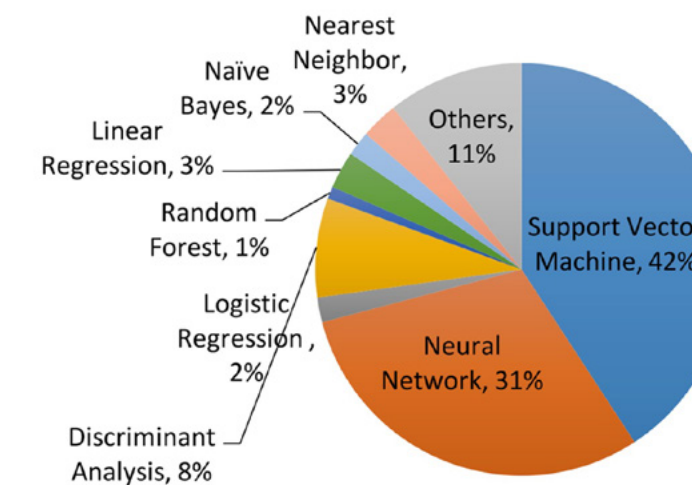
- In letteratura spesso alcune *features* sono interpretate come biomarcatori specifici ad es. di una patologia, di un sottotipo molecolare ecc.
- due approcci distinti e alternativi:
 1. priorità all'**ipotesi e interpretazione** => ricerca di modelli predittivi basati su poche (o singole) *feature* selezionate
 2. priorità alle **prestazioni** => ricerca di modelli predittivi il più performanti possibili => approccio «agnostico», senza assunzioni a priori

Radiomica e intelligenza artificiale

- Grande insieme di dati non omogenei
- Complessità tecnica
- **Overfitting**
- Adeguatezza del **disegno sperimentale**
- **Validazione** (interna o esterna)
- **Confounding variables**
- **Interpretazione** dei risultati



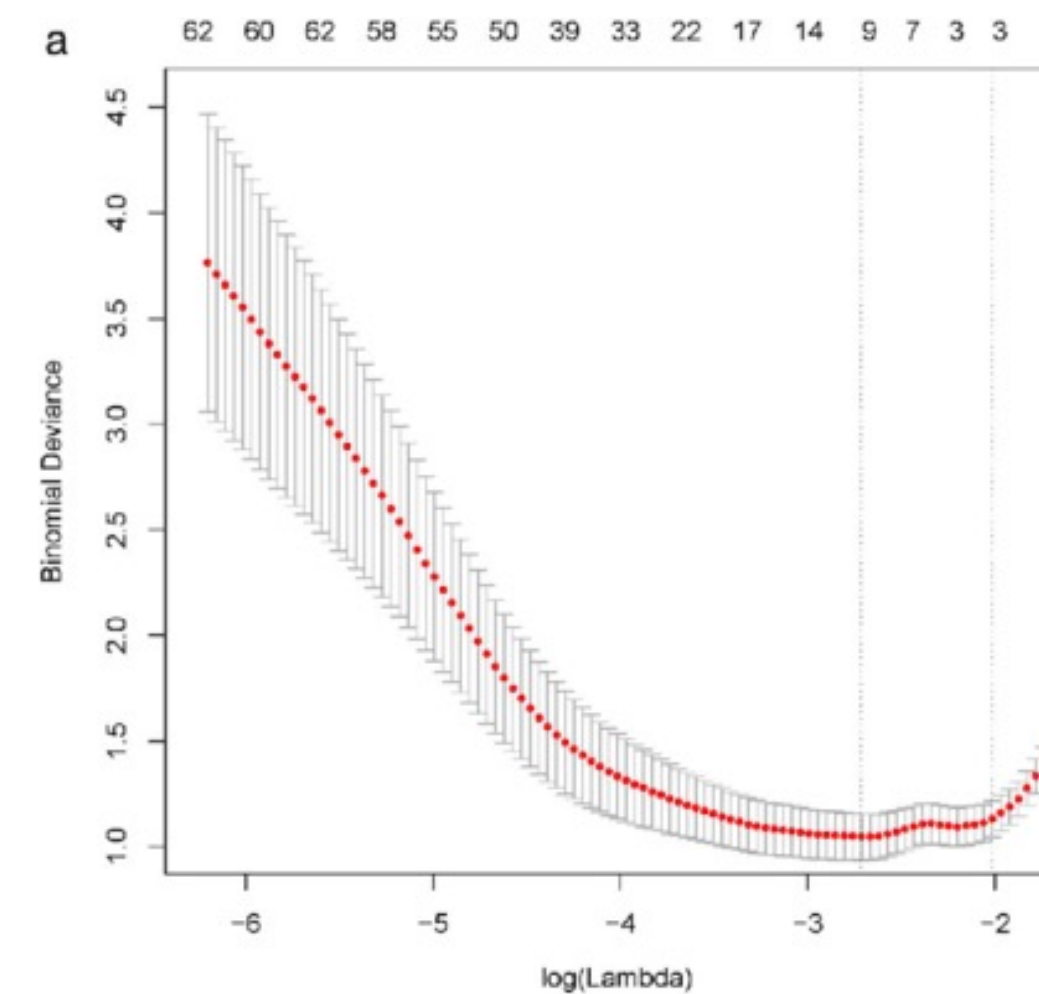
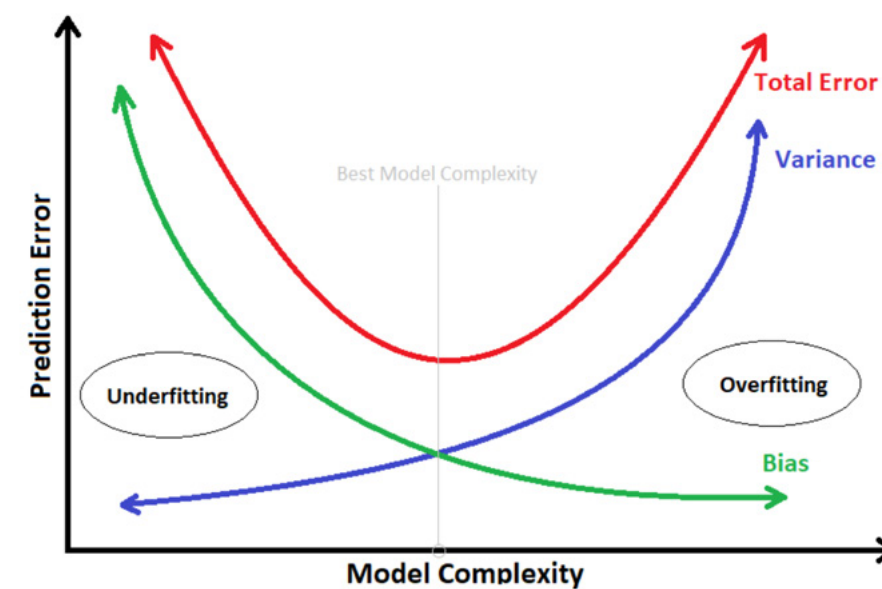
machine learning
per estrazione di
informazione



Jiang F, Jiang Y, Zhi H, et al. Artificial intelligence in healthcare: past, present and future. *Stroke and Vascular Neurology* 2017;2: e000101. doi:10.1136/svn-2017-000101

Modelli ad alte prestazioni

- modelli complessi (elevato numero di parametri) => maggiore accuratezza (low bias) ma esposizione a overfitting (high variance)



#sanita2030



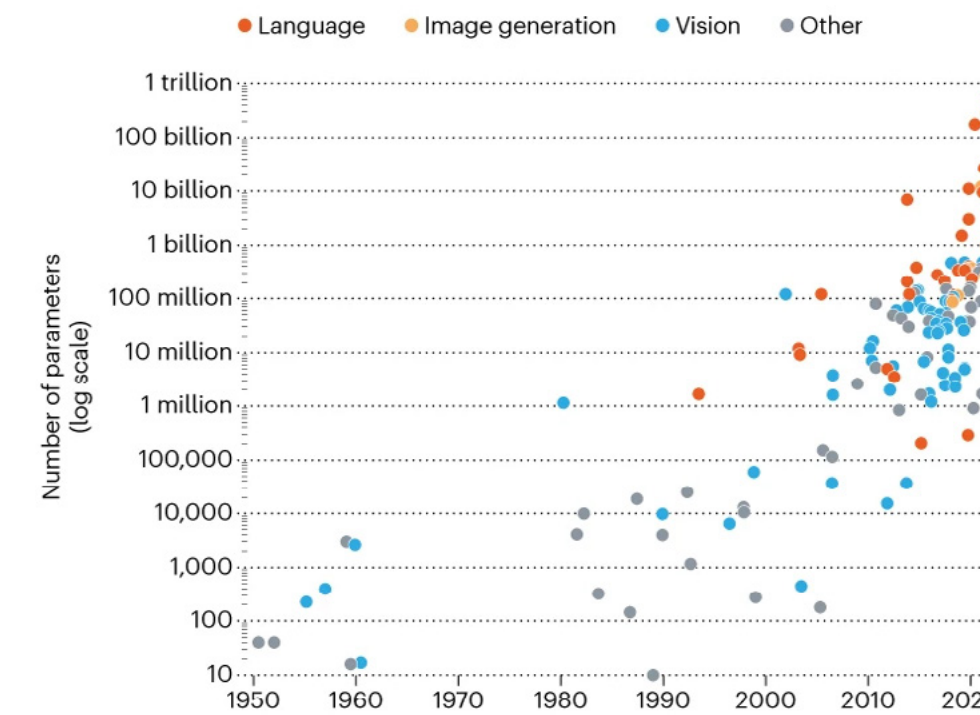
www.sanita2030.it



Modelli ad alte prestazioni – *foundation models*

- "models trained on broad data (generally using self-supervision at scale) that can be adapted to a wide range of downstream tasks"*
- modelli **generalizzabili** (general-purpose AI - **GPAI**)
- **architettura comune** ad elevate prestazioni che viene adattata (**fine tuning**) a problemi specifici modificando solo una piccola parte del modello (vs. narrow-AI)
- **altissimo numero di parametri**
- richiedono training su database di enormi dimensioni => **costi elevati** (1-10 M\$)

* Stanford Institute for Human-Centered Artificial Intelligence, 'Reflections on Foundation Models' <<https://hai.stanford.edu/news/reflections-foundation-models>>, accessed 5 November 2023.

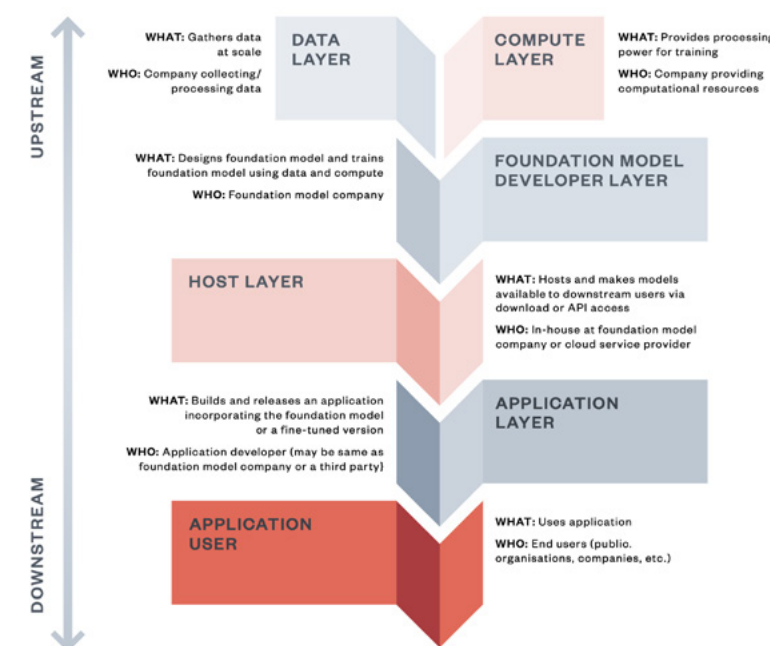


Nature, news section, 14 March 2023

Adapted from: J. Sevilla et al. Preprint at <https://arxiv.org/abs/2202.05924> (2022)

Modelli ad alte prestazioni – *foundation models*

Foundation model supply chain



Note: This is one possible model (there will not always be a separate or single company at each layer)

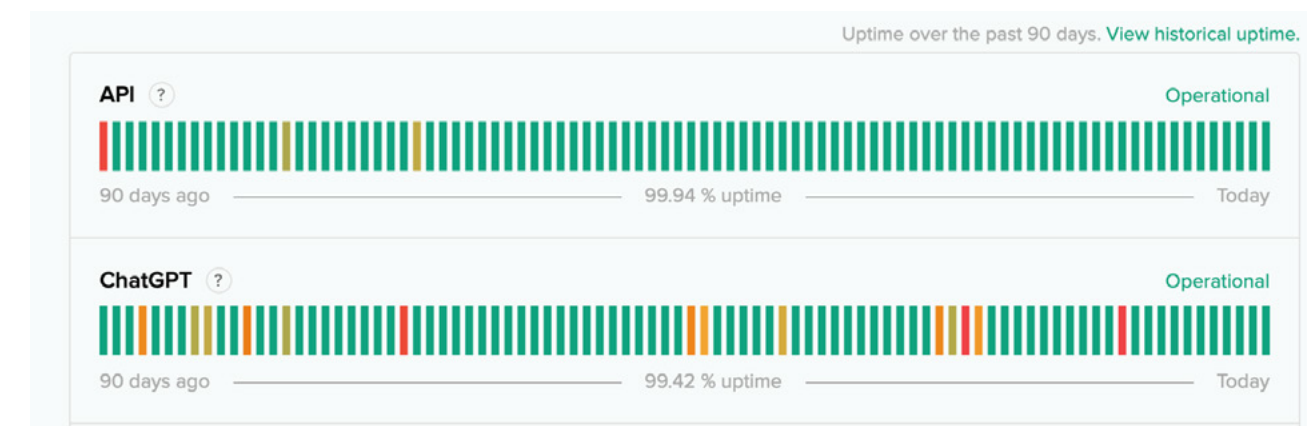
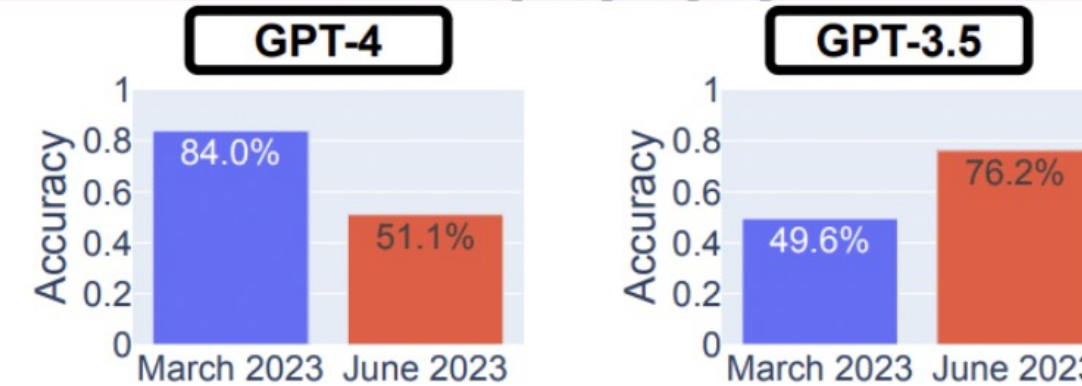
<https://www.adalovelaceinstitute.org/resource/foundation-models-explainer/>
accessed 12 June 2024

- Sviluppo possibile solo in un "ecosistema" complesso
- **Database** per training e **risorse di calcolo** inaccessibili a singoli gruppi di ricerca e sviluppo (al ritmo attuale raddoppio delle capacità di calcolo ogni 6-10 mesi)
- Necessità di **regolamentazione e controllo**
 - sicurezza
 - privacy
 - **QA** – garanzia di **mantenimento delle prestazioni**

Qualità e sicurezza

- processo **non completamente tracciabile** per
 - complessità
 - evoluzione continua (life-long learning)
- possibile **degrado di prestazioni** imprevedibile a causa dell'adattamento continuo ai dati
- necessario garantire **robustezza degli algoritmi**

Is 17077 a prime number? Think step by step and then answer [Yes] or [No].



<https://status.openai.com>

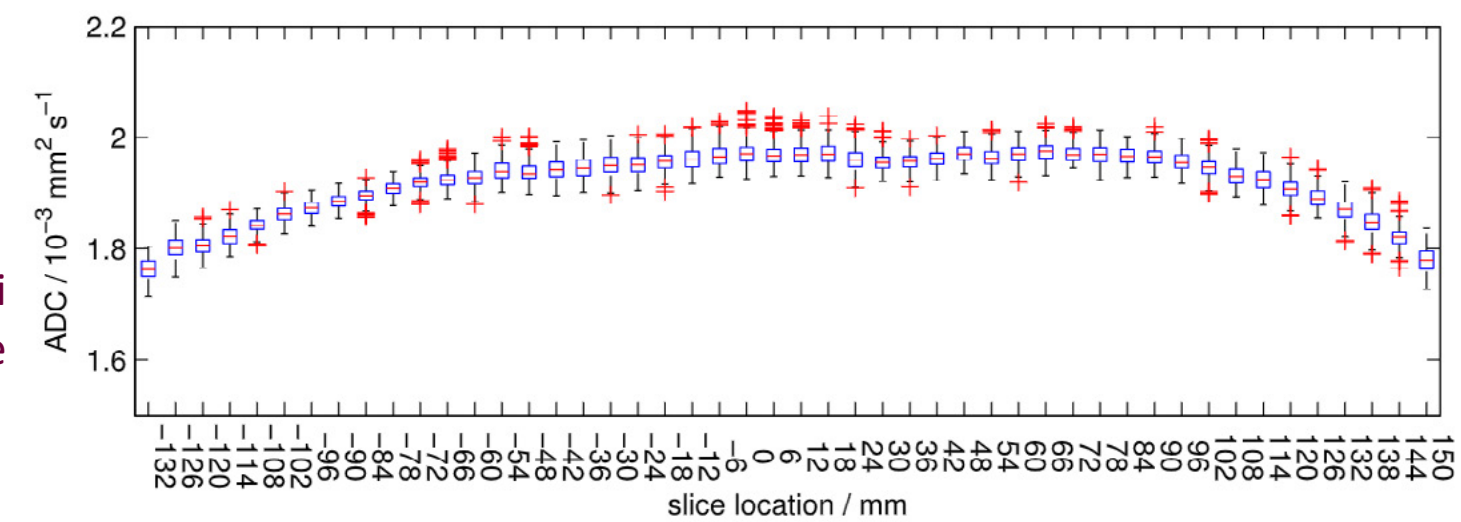
#sanita2030



www.sanita2030.it

Qualità e sicurezza: la qualità dei dati in ingresso

- *standardizzazione*
 - dati omogenei in partenza
 - *omogeneizzazione*
 - dati resi confrontabili
- ESEMPIO:
- variabilità inter-scanner in MRI
 - influenza dei parametri di acquisizione (ad es. posizione della slice) sui dati quantitativi



JM Winfield, DJ Collins, AN Priest et al. "A framework for optimization of diffusion-weighted MRI protocols for large field-of-view abdominal-pelvic imaging in multicenter studies". Med. Phys. 43(1), pp. 95-110 (2016).

Qualità e sicurezza

- **sviluppo di algoritmi ad alte prestazioni** riservato a pochi player internazionali
- **sicurezza e qualità** su tutti i piani, compreso l'utilizzatore finale
- **facilitazione e semplificazione**
- programma di garanzia della qualità con **controlli strutturati**
 - dati in ingresso
 - mantenimento prestazioni

GOV.UK

Home > Business and Industry > Science and Innovation > Artificial Intelligence > AI Safety Summit 2023: The Bletchley Declaration

Department for Science, Innovation & Technology | Foreign, Commonwealth & Development Office | Prime Minister's Office, 10 Downing Street

Policy paper
The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023
Published 1 November 2023

In the context of our cooperation, and to inform action at the national and international levels, our agenda for addressing frontier AI risk will focus on:

- identifying AI safety risks of shared concern, building a shared scientific and **evidence-based understanding of these risks**, and sustaining that understanding as capabilities continue to increase, in the context of a wider global approach to understanding the impact of AI in our societies.
- building respective risk-based policies across our countries to ensure safety in light of such risks, collaborating as appropriate while recognising our approaches may differ based on national circumstances and applicable legal frameworks. This includes, alongside **increased transparency** by private actors developing frontier AI capabilities, **appropriate evaluation metrics**, tools for **safety testing**, and developing relevant public sector capability and scientific research.

#sanita2030



www.sanita2030.it





Grazie

carlo.cavedon@aovr.veneto.it

Delitti in materia di violazione del diritto d'autore (Art. 25-novies, D.Lgs. n. 231/2001) [articolo aggiunto dalla L. n. 99/2009]

- Messa a disposizione del pubblico, in un sistema di reti telematiche, mediante connessioni di qualsiasi genere, di un'opera dell'ingegno protetta, o di parte di essa (art. 171, legge n.633/1941 comma 1 lett. a) bis)
- Reati di cui al punto precedente commessi su opere altrui non destinate alla pubblicazione qualora ne risulti offeso l'onore o la reputazione (art. 171, legge n.633/1941 comma 3)
- Abusiva duplicazione, per trarne profitto, di programmi per elaboratore; importazione, distribuzione, vendita o detenzione a scopo commerciale o imprenditoriale o concessione in locazione di programmi contenuti in supporti non contrassegnati dalla SIAE; predisposizione di mezzi per rimuovere o eludere i dispositivi di protezione di programmi per elaboratori (art. 171-bis legge n.633/1941 comma 1)
- Riproduzione, trasferimento su altro supporto, distribuzione, comunicazione, presentazione o dimostrazione in pubblico, del contenuto di una banca dati; estrazione o reimpiego della banca dati; distribuzione, vendita o concessione in locazione di banche di dati (art. 171-bis legge n.633/1941 comma 2)
- Abusiva duplicazione, riproduzione, trasmissione o diffusione in pubblico con qualsiasi procedimento, in tutto o in parte, di opere dell'ingegno destinate al circuito televisivo, cinematografico, della vendita o del noleggio di dischi, nastri o supporti analoghi o ogni altro supporto contenente fonogrammi o videogrammi di opere musicali, cinematografiche o audiovisive assimilate o sequenze di immagini in movimento; opere letterarie, drammatiche, scientifiche o didattiche, musicali o drammatico musicali, multimediali, anche se inserite in opere collettive o composite o banche dati; riproduzione, duplicazione, trasmissione o diffusione abusiva, vendita o commercio, cessione a qualsiasi titolo o importazione abusiva di oltre cinquanta copie o esemplari di opere tutelate dal diritto d'autore e da diritti connessi; immissione in un sistema di reti telematiche, mediante connessioni di qualsiasi genere, di un'opera dell'ingegno protetta dal diritto d'autore, o parte di essa (art. 171-ter legge n.633/1941)
- Mancata comunicazione alla SIAE dei dati di identificazione dei supporti non soggetti al contrassegno o falsa dichiarazione (art. 171-septies legge n.633/1941)
- Fraudolenta produzione, vendita, importazione, promozione, installazione, modifica, utilizzo per uso pubblico e privato di apparati o parti di apparati atti alla decodificazione di trasmissioni audiovisive ad accesso condizionato effettuate via etere, via satellite, via cavo, in forma sia analogica sia digitale (art. 171-octies legge n.633/1941).

[Torna all'inizio](#)