

Intelligenza artificiale nella gestione del paziente settico: luci e ombre...

Germana Ruggiano
SOC Medicina d'Urgenza
USL Toscana Centro



Punti chiave

- Concetti generali
- Applicazioni nei pazienti con sepsi in Medicina d'urgenza
- Aspetti controversi



- **Concetti generali**
- Applicazioni nei pazienti con sepsi in Medicina d'urgenza
- Aspetti controversi



Introduction to Machine Learning, Neural Networks, and Deep Learning

- Nel 1956, un gruppo di informatici aveva proposto che i computer potevano essere programmati per pensare e ragionare, che “ogni aspetto dell’apprendimento e tutte le altre manifestazioni dell’intelligenza potrebbero in principio essere descritte così’ precisamente che potrebbe essere costruita una macchina per simularle. Descrissero questo principio con “artificial intelligence.”

1. Brown JM, Campbell JP, Beers A, et al. Automated diagnosis of plus disease in retinopathy of prematurity using deep convolutional neural networks. *JAMA Ophthalmol.* 2018;136:803–810. doi:10.1001/jamaophthalmol.2018.1934.

2. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA.* 2016;316:2402–2410. doi:10.1001/jama.2016.17216.

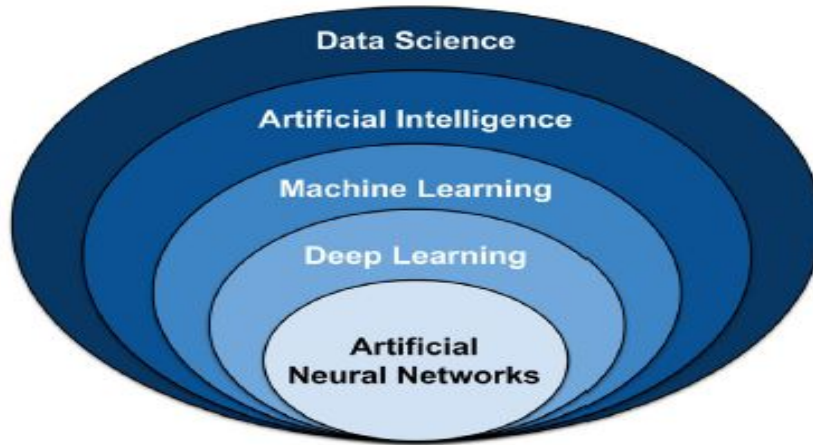
3. Coyner AS, Swan R, Campbell JP, et al. Automated fundus image quality assessment in retinopathy of prematurity using deep convolutional neural networks. *Ophthalmol Retina.* 2019;3:444–450. doi:10.1016/j.oret.2019.01.015.

4. Rajpurkar P, Irvin J, Zhu K, et al. CheXNet: radiologist-level pneumonia detection on chest Xrays with deep learning. *ArXiv171105225 Cs Stat.* November 2017. <http://arxiv.org/abs/1711.05225>. Accessed October 23, 2019.

5. Jones LD, Golan D, Hanna SA, Ramachandran M. Artificial intelligence, machine learning and the evolution of healthcare: a bright future or cause for concern? *Bone Jt Res.* 2018;7:223–225. doi:10.1302/2046-3758.73.BJR-2017-0147.R1.

6. De Fauw J, Ledsam JR, Romera-Paredes B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med.* 2018;24:1342–1350. doi:10.1038/s41591-018-0107-6.

- AI e' focalizzata sull'automatizzazione di compiti normalmente svolti da parte degli uomini, e ML e DL sono metodiche specifiche per raggiungere questi obiettivi ; quindi sono all'interno della AI



Anche se l'AI e' eccellente nel risolvere problemi di logica chiaramente definiti, spesso fallisce in compiti che richiedano livelli piu' alti di riconoscimento di pattern come il riconoscimento di discorsi o la classificazione di immagini. Questi compiti piu' complessi sono quelli in cui ML e DL riescono meglio.

Machine Learning: si focalizza sugli aspetti di apprendimento dell'AI sviluppando algoritmi che rappresentano al meglio un set di dati

- In ML, ci sono 4 metodi comuni di apprendimento ognuno utile per compiti differenti: supervised , unsupervised, semisupervised e “reinforcement learning”.
- Supervised learning: all'algoritmo vengono dati dati con un output noto ed il modello e' sviluppato per predire quell'output a partire da diversi input: i piu' comuni sono la regressione e la classificazione.
- La Regressione include la previsione a partire da dati numerici come gli scores prognostici, valori di laboratorio, prezzi di un prodotto.
- La classificazione, prevede l'individuazione della categoria a cui appartiene un dato esempio.

simeu

James G, Witten D, Hastie T, Tibshirani R, eds. *An Introduction to Statistical Learning: With Applications in R*. New York: Springer; 2013.

Hastie T, Tibshirani R, Friedman JH. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. New York, NY: Springer; 2009.

NVIDIA Blog: Supervised Vs. Unsupervised Learning. The Official NVIDIA Blog. <https://blogs.nvidia.com/blog/2018/08/02/supervisedunsupervised-learning/>. Published August 2, 2018.

Unsupervised learning

- L'apprendimento non supervisionato ha come obiettivo di evidenziare patterns in un set di dati e caratterizzare casi individuali nel data set di una certa categoria.
- Questi algoritmi sono non supervisionati perché i pattern in un data set non sono costruiti a partire da un output già noto e sono lasciati alla determinazione da parte dell'algoritmo

Semisupervised learning

- L'apprendimento semisupervisionato può essere considerato come il “giusto mezzo” tra l'apprendimento supervisionato e quello non supervisionato ed è particolarmente utile per database che contengono sia dati catalogati che non catalogati



XIII congresso nazionale

simeu

James G, Witten D, Hastie T, Tibshirani R, eds. *An Introduction to Statistical Learning: With Applications in R*. New York: Springer; 2013.

Hastie T, Tibshirani R, Friedman JH. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. New York, NY: Springer; 2009.

NVIDIA Blog: Supervised Vs. Unsupervised Learning. The Official NVIDIA Blog. <https://blogs.nvidia.com/blog/2018/08/02/supervisedunsupervised-learning/>. Published August 2, 2018.

GENOVA 30 MAG - 1 GIU 2024

“Reinforcement Learning”

- “reinforcement learning” e’ la tecnica di “allenamento” di un algoritmo per un compito specifico dove non c’e’ un’unica risposta giusta ma si desidera un certo outcome.
- E’ il tentativo piu’ vicino al modello di apprendimento dell’uomo perche’ impara anche dai tentativi e dagli errori oltre che dai soli dati.
- Anche se “reinforcement” ha un certo rilievo all’interno della computer science e machine learning, al momento non ha un impatto sostanziale nella medicina clinica.

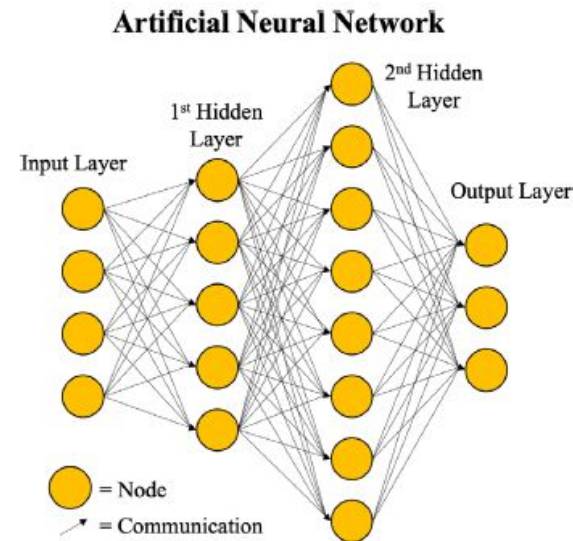


Hastie T, Tibshirani R, Friedman JH. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. New York, NY: Springer; 2009.

NVIDIA Blog: Supervised Vs. Unsupervised Learning. The Official NVIDIA Blog. <https://blogs.nvidia.com/blog/2018/08/02/supervisedunsupervised-learning/>. Published August 2, 2018. Accessed October 24, 2019.

Neural Networks e Deep Learning

- Un artificial neural network (ANN) e' un algoritmo di machine learning ispirato al network neurale biologico. Ogni ANN contiene nodi (analoghi ai corpi cellulari) che comunicano con altri nodi attraverso connessioni (analoghi agli assoni ed ai dendriti).
- Così' come le sinapsi tra i neuroni sono rinforzate quando i neuroni hanno outputs correlati (la teoria Hebbiana postula che "nerves that fire together, wire together"), le connessioni tra i nodi di un ANN sono "pesate" sulla loro abilità di fornire un outcome desiderato.



Come viene costruito lo studio con ML

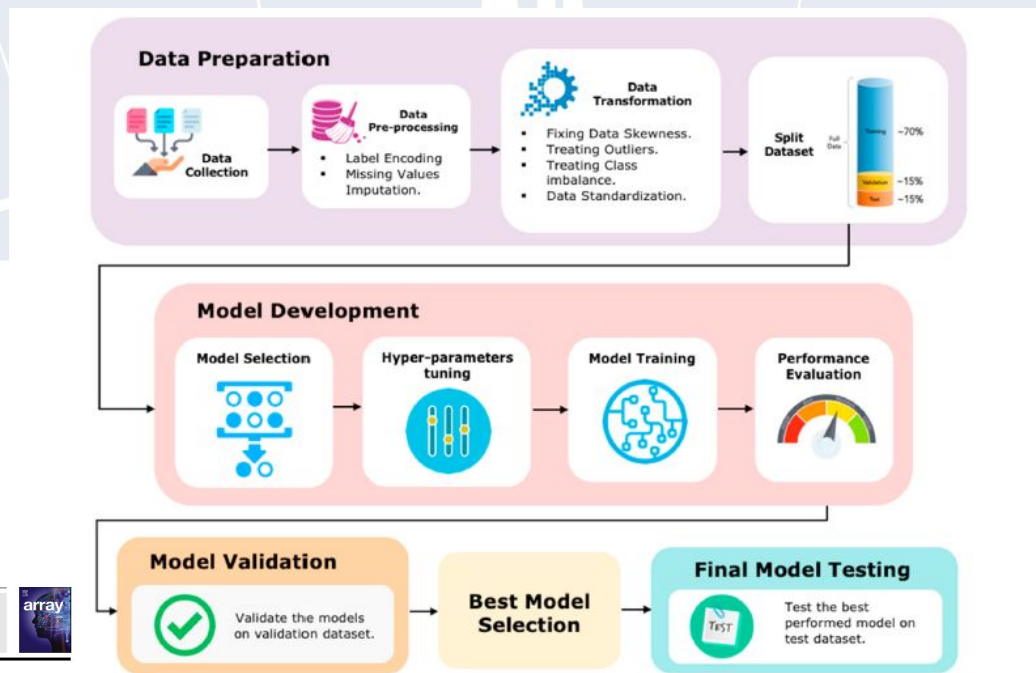


Fig. 1. Study experimental setup.

Applicazioni della ML e della medicina di precisione alla ricerca in EM

1) AI nell'imaging

L'uso della AI nell'imaging puo' aiutare i clinici a fare diagnosi tempestive processando le immagini per ottimizzare la visualizzazione, evidenziando le alterazioni

Gran parte della ricerca attuale sulla AI nell'imaging si focalizza sull'uso del deep learning (DL) per assistere il radiologo nel riconoscimento di pattern particolari o misurazioni quantitative

Fuori dal campo radiologico e' stata condotta ricerca attiva in AI anche nell'ecocardiografia.





2) Processi di Natural language (NLP) nel setting della EM

- NLP e' un gruppo di AI dedicato all'analisi ed al processo dei dati del linguaggio umano
- Utilizza algoritmi di AI per processare testi o audio da sorgenti non strutturate, estrarre informazioni e convertirle in dati strutturati processabili dai computer.
- E' stata utilizzata per processare dati provenienti da EHR, per esempio con il riconoscimento della voce per la registrazione dei dati
- Attraverso l'analisi delle informazioni registrate nelle EHR, l'approccio con NLP puo' combinare multiple sorgenti di dati ed aiutare i professionisti nella diagnostica di molte patologie, nella valutazione

dell'appropriatezza della gestione dei traumi della strada, nella capacita' di predire l'ammissione in

Ohno-Machado L, Nadkarni P, Johnson K: Natural language processing: algorithms and tools to extract computable information from EHRs and from the biomedical literature. *J Am Med Inform Assoc.* 2013, 20:805. 10.1136/amiajnl-2013-002214

Doan S, Maehara CK, Chaparro JD, et al.: Building a natural language processing tool to identify patients with high clinical suspicion for Kawasaki disease from emergency department notes. *Acad Emerg Med.* 2016, 23:628-36. 10.1111/acem.12925

Tignanelli CJ, Silverman GM, Lindemann EA, et al.: Natural language processing of prehospital emergency medical services trauma records allows for automated characterization of treatment appropriateness. *J Trauma Acute Care Surg.* 2020, 88:607-14. 10.1097/TA.0000000000002598

Sterling NW, Patzer RE, Di M, Schrager JD: Prediction of emergency department patient disposition based on natural language processing of triage notes. *Int J Med Inform.* 2019, 129:184-8. 10.1016/j.ijmedinf.2019.06.008

Klang E, Kummer BR, Dangayach NS, et al.: Predicting adult neuroscience intensive care unit admission from emergency department triage using a retrospective, tabular-free text machine learning approach. *Sci Rep.* 2021, 11:1381. 10.1038/s41598-021-80985-3

3) Ricerca nella Precision medicine

- “Precision medicine” consiste nell’assicurare il trattamento piu’ adeguato ad un determinato paziente.
- Il concetto di base e’ che la diagnosi di per se non e’ sufficiente per stabilire il trattamento adeguato che deve essere “fatto su misura” partendo da dati individuali del paziente. Spesso questo include una fenotipizzazione molecolare sistematica (-omics).
- La combinazione di metabolomics con le variabili cliniche crea predittori piu’ accurati di mortalita’ per sepsi rispetto alle sole varabili cliniche o emodinamiche

Collins FS, Varmus H: A new initiative on precision medicine . *N Engl J Med*. 2015, 372:793-5. 10.1056/NEJMp1500523

Zaas AK, Burke T, Chen M, et al.: A host-based RT-PCR gene expression signature to identify acute respiratory viral infection. *Sci Transl Med*. 2013, 5:203ra126. 10.1126/scitranslmed.3006280

Finnerty CC, Jeschke MG, Qian WJ, et al.: Determination of burn patient outcome by large-scale quantitative discovery proteomics. *Crit Care Med*. 2013, 41:1421-34. 10.1097/CCM.0b013e31827c072e

Warren HS, Elson CM, Hayden DL, et al.: A genomic score prognostic of outcome in trauma patients . *Mol Med*. 2009, 15:220-7. 10.2119/molmed.2009.00027

Wong HR, Salisbury S, Xiao Q, et al.: The pediatric sepsis biomarker risk model. *Crit Care*. 2012, 16:R174. 10.1186/cc11652

Biffi A, Anderson CD, Jagiella JM, et al.: APOE genotype and extent of bleeding and outcome in lobar intracerebral haemorrhage: a genetic association study. *Lancet Neurol*. 2011, 10:702-9. 10.1016/S1474-4422(11)70148-X

Monte AA, Heard KJ, Campbell J, Hamamura D, Weinshilboum RM, Vasiliou V: The effect of CYP2D6 drugdrug interactions on hydrocodone effectiveness. *Acad Emerg Med*. 2014, 21:879-85. 10.1111/acem.12431

Puskarich MA, Finkel MA, Karnovsky A, Jones AE, Trexel J, Harris BN, Stringer KA: Pharmacometabolomics of l-carnitine treatment response phenotypes in patients with septic shock. *Ann Am Thorac Soc*. 2015, 12:46-56. 10.1513/AnnalsATS.201409-415OC

del Río-Espínola A, Fernández-Cadenas I, Giralte D, et al.: A predictive clinical-genetic model of tissue plasminogen activator response in acute ischemic stroke. *Ann Neurol*. 2012, 72:716-29. 10.1002/ana.23664

Langley RJ, Tsalik EL, van Velkinburgh JC, et al.: An integrated clinico-metabolomic model improves prediction of death in sepsis. *Sci Transl Med*. 2013, 5:195ra95. 10.1126/scitranslmed.3005893

- 
- Concetti generali
 - Applicazioni nei pazienti con sepsi in Medicina d'urgenza
 - Aspetti controversi



Review

Jie Yang, Sicheng Hao, Jiajie Huang, Tianqi Chen, Ruoyi Liu, Ping Zhang, Mengling Feng, Yang He, Wei Xiao, Yucai Hong and Zhongheng Zhang*

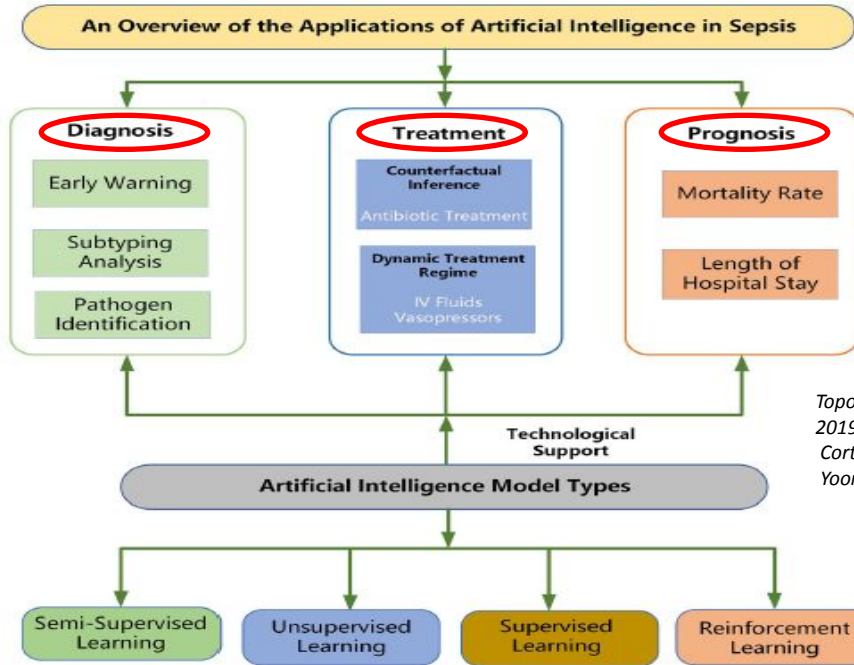
The application of artificial intelligence in the management of sepsis

- Nonostante l'evoluzione della Critical care medicine, la sepsi resta la causa principale di morte nei pazienti critici.
- La complessita' nella diagnosi e nel trattamento della sepsi sono legate all'eterogeneita' della patologia.
- Pazienti settici presentano manifestazioni cliniche differenti e rispondono in modo differente ad un particolare trattamento
- Richiedono spesso una gestione altamente personalizzata "precision medicine".

Review

Jie Yang, Sicheng Hao, Jiajie Huang, Tianqi Chen, Ruoqi Liu, Ping Zhang, Mengling Feng, Yang He, Wei Xiao, Yucai Hong and Zhongheng Zhang*

The application of artificial intelligence in the management of sepsis



L'integrazione della (AI) nel dominio della medicina sta assumendo un ruolo nella

- diagnosi,
- prognosi e
- trattamento di varie patologie critiche inclusa la sepsi

Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med* 2019;25:44–56.

Corte TD, Hoecke SV, Waele JD. Artificial intelligence in infection management in the ICU. *Crit Care* 2022;26:79.

Yoon JH, Pinsky MR, Clermont G. Artificial intelligence in critical care medicine. *Crit Care* 2022;26:75.

Figure 2: An overview of the application of artificial intelligence in sepsis. IV fluids, intravenous fluids.

Jie Yang, Sicheng Hao, Jiajie Huang, Tianqi Chen, Ruoli Liu, Ping Zhang, Mengling Feng, Yang He, Wei Xiao, Yucui Hong and Zhongheng Zhang*

The application of artificial intelligence in the management of sepsis

Table 3: Publicly accessible intensive care databases.

Name	Data collection timeframe	Patients volume	Country	Patient sources	Website
MIMIC-IV	2008–2019	299,712	The United States	Admitted to intensive care units at the beth Israel deaconess medical center (BIDMC)	https://physionet.org/content/mimiciv/2.2/
eICU	2014–2015	Over 200,000	The United States	Admitted to one of 335 units at 208 hospitals located throughout the US	https://physionet.org/content/eicu-crd/2.0/
AmsterdamUMCdb	2003–2016	20,109	The Netherlands	Admitted to intensive care units at an academic medical center in Amsterdam	https://github.com/AmsterdamUMC/AmsterdamUMCdb
HiRID	2008–2016	34,000	Switzerland	Admissions to the department of intensive care medicine of the bern University hospital	https://physionet.org/content/hirid/1.1.1/
PIC	2010–2018	12,881	China	Paediatric patients (aged 0–18 years) admitted to critical care unit	https://physionet.org/content/picdb/1.1.0/
ZhejiangProvinceICU	2012–2022	7,638	China	Admitted to intensive care units and emergency intensive care units at Zhejiang provincial People's hospital	https://www.nature.com/articles/s41597-023-01952-3
INSPIRE	2011–2020	130,000	South Korea	Underwent anesthesia for surgery at an academic institution	https://www.physionet.org/content/inspire/0.1/

Esempi di database pubblici che vengono utilizzati per la machine learning

1) Diagnosi di sepsi

“Early warning”: multipli modelli di AI sono stati progettati per prevedere lo sviluppo di sepsi

- Attraverso il monitoraggio continuo di dati clinici, l'esordio della sepsi può essere previsto numerose ore prima con un'accuratezza vicina al 90 %, miglioramento sostanziale rispetto a score tradizionali
- Metanalisi recenti hanno rivelato che i benefici dei sistemi di early warning con AI sono maggiori nei dipartimenti di emergenza e nei reparti di degenza rispetto alle ICU

Islam MM, Nasrin T, Walther BA, Wu CC, Yang HC, Li YC. Prediction of sepsis patients using machine learning approach: a meta-analysis. Comput Methods Progr Biomed 2019;170:1–9.

Zhang Z, Chen L, Xu P, Wang Q, Zhang J, Chen K, et al. Effectiveness of automated alerting system compared to usual care for the management of sepsis. NPJ Digit Med 2022;5:101.

simeu

GENOVA 30 MAG - 1 GIU 2024

Table 3
Performance of predicting sepsis by commonly used methods.

Name	No. study	AUROC	Sensitivity (95%CI)	Specificity (95%CI)	PLR(95%CI)	NLR(95%CI)	Diagnostic odd ratio
ML	7	0.89	0.81 (0.80-0.81)	0.72 (0.72-0.72)	3.31(1.46-7.48)	0.23(0.20-0.27)	15.17 (9.51-24.20)
SIRS	2	0.70	0.75 (0.74-0.77)	0.50 (0.50-0.50)	1.50(1.19-1.88)	0.46(0.27-0.80)	3.23 (1.52-6.87)
MEWS	2	0.50	0.80 (0.79-0.81)	0.73 (0.73-0.73)	3.30(2.69-4.04)	0.10(0.00-2.43)	31.99 (1.54-666.77)
SOFA	3	0.78	0.77 (0.76-0.78)	0.42 (0.42-0.42)	1.68(1.13-2.52)	0.45 (0.37-0.54)	3.75 (2.06-6.83)
QSOFA	1	0.77	0.56 (0.54-0.57)	0.84 (0.83-0.84)	3.50 (3.34-3.67)	0.52 (0.50-0.55)	N/A

*Note: ML=Machine learning, PLR=Positive Likelihood Ratio, NLR=Negative Likelihood Ratio, SIRS=Systemic inflammatory response syndrome, MEWS=Modified early warning system, SOFA=Sequential organ failure assessment, QSOFA=Quick sequential organ failure assessment.

La performance della ML per il riconoscimento precoce della sepsi e' migliore rispetto a strumenti classici come SIRS, MEWS, SOFA, e QSOFA.

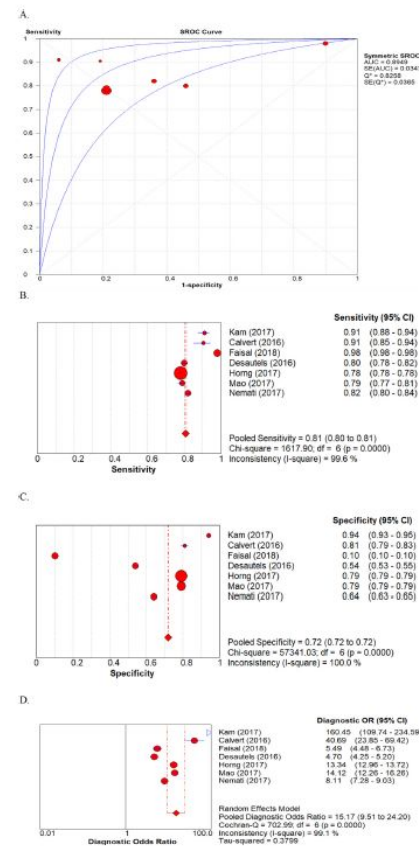


Fig. 2. Performance of machine learning model for predicting sepsis (A. AUROC, B. Sensitivity C. Specificity D. Diagnostic odd ratio). The color circle represents proportion of the number of patients with sepsis and the number of patients without sepsis. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2) Sottotipizzazione della sepsi

L'utilizzo di algoritmi che permettano di **distinguere diversi sottotipi di sepsi** e' un focus di ricerca recente .

- Analisi retrospettive di variabili cliniche utilizzando ML possono classificare la sepsi in **4 fenotipi: α , β , γ , and δ** . Questi sottotipi rappresentano profili clinici differenti o manifestazioni differenti ognuno con le sue caratteristiche e outcome unici
- Allo stesso modo il **danno renale** correlato alla sepsi e' stato suddiviso in 3 sottotipi distinti basandosi su markers biochimici, pattern immunologici e outcomes clinici.
- I ricercatori hanno anche classificato la sepsi in 2 sottotipi basati su **differenze nell'espressione di alcuni geni** attraverso analisi estese di dati. Questi sottotipi presentano differenze significative di mortalita' e risposta a terapia ormonale

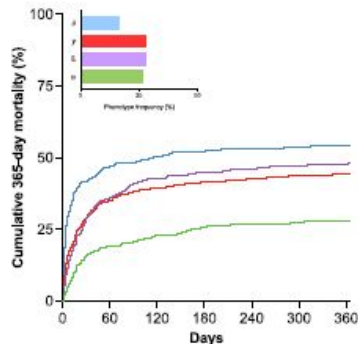
Seymour CW, Kennedy JN, Wang S, Chang CH, Elliott CF, Xu Z, et al. Derivation, validation, and potential treatment implications of novel clinical phenotypes for sepsis. JAMA 2019;321:2003.

Chaudhary K, Vaid A, Duffy Á, Paranjpe I, Jaladanki S, Paranjpe M, et al. Utilization of deep learning for subphenotype identification in sepsis-associated acute kidney injury. Clin J Am Soc Nephrol 2020;15: 1557–65.

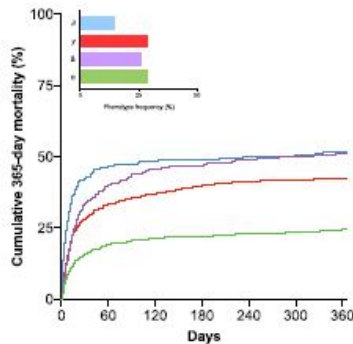
Zhang Z, Pan Q, Ge H, Xing L, Hong Y, Chen P. Deep learning-based clustering robustly identified two classes of sepsis with both prognostic and predictive values. EBioMedicine 2020;62:103081

Derivation, Validation, and Potential Treatment Implications of Novel Clinical Phenotypes for Sepsis

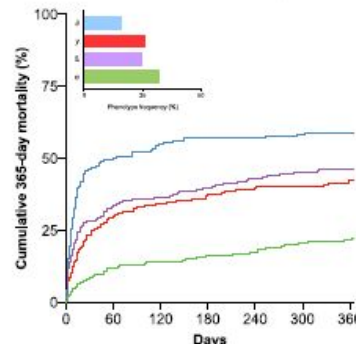
A ACCESS trial (N=1,706)
(Eritoran vs. placebo)



B PROWESS trial (N=1,690)
(Drotrecogin alfa vs. placebo)



C ProCESS trial (N=1,341)
(protocolized EGDT vs. protocolized standard care vs. usual care)



4 fenotipi: α , β , γ ,
and δ con
manifestazioni
differenti e
outcome unici

(A) 365-day mortality in the ACCESS trial (N=1,706), (B) 365-day mortality in PROWESS trial (N=1,690), and (C) 365-day mortality in the ProCESS trial, by phenotype, where α -type is light green, β -type is light purple, γ -type is light red, and δ -type is light blue. All panels show significant differences in mortality by phenotype (log rank $P < 0.001$). Inset histograms show the proportion of patients in each phenotype.

Interpretive example: In 3 randomized clinical trials with follow up to 365 days after enrollment, phenotypes are significantly and consistently associated with outcome.

3) Identificazione dei patogeni e test della suscettibilità' antimicrobica

- La somministrazione precoce degli antibiotici in emergenza e' su base empirica Per migliorare l'efficacia della diagnostica microbiologica sono stati sviluppati vari algoritmi di AI .
- Questi modelli **acquisiscono dati da differenti fonti incluso il microbiota, genomics, sequenze geniche, morfologia di colonia, immagini da microscopio, transcriptomics, e dati clinici**
- Alcuni ricercatori hanno combinato l'analisi di campioni di sangue con modelli di AI per identificare rapidamente 30 batteri e funghi comuni insieme a suscettibilità' con **un'accuratezza di oltre l' 80%**
- Sono stati sviluppati modelli di AI per **prevedere la resistenza agli antimicrobici** e ottimizzare la terapia antibiotica empirica



Peiffer-Smadja N, Dellièvre S, Rodriguez C, Birgand G, Lescure FX, Fourati S, et al. Machine learning in the clinical microbiology laboratory: has the time come for routine practice? Clin Microbiol Infect 2020;26:1300–9.

Ho CS, Jean N, Hogan CA, Blackmon L, Jeffrey SS, Holodniy M, et al. Rapid identification of pathogenic bacteria using Raman spectroscopy and deep learning. Nat Commun 2019;10:4927.

Feretzakis G, Loupelis E, Sakagianni A, Kalles D, Martsoukou M, Lada M, et al. Using machine learning techniques to aid empirical antibiotic therapy decisions in the intensive care unit of a general hospital in Greece. Antibiotics 2020;9:50.

Moran E, Robinson E, Green C, Keeling M, Collyer B. Towards personalized guidelines: using machine-learning algorithms to guide antimicrobial selection. J Antimicrob Chemother 2020;75: 2677–80.

Roimi M, Neuberger A, Shrot A, Paul M, Geffen Y, Bar-Lavie Y. Early diagnosis of bloodstream infections in the intensive care unit using machine-learning algorithms. Intensive Care Med 2020;46:454–62.

Steenkiste TV, Ruysinck J, Baets LD, Decruyenaere J, Turck FD, Ongenaes F, et al. Accurate prediction of blood culture outcome in the intensive care unit using long short-term memory neural networks. Artif Intell Med 2019;97:38–43.

Garnica O, Gómez D, Ramos V, Hidalgo JI, Ruiz-Giardin JM. Diagnosing hospital bacteraemia in the framework of predictive, preventive and personalised medicine using electronic health records and machine learning classifiers. EPMA J 2021;12:365–81.

4) Trattamento di precisione della sepsi

- La somministrazione precoce di fluidi puo' migliorare la perfusione dei tessuti aumentando il ritorno venoso e il cardiac output ma portare a edema significativo con disfunzione d'organo
- I vasopressori possono essere utilizzati per risolvere l'ipotensione e mantenere la perfusione limitando la somministrazione di fluidi.
- Nella gestione della sepsi possono essere utilizzati modelli di reinforcement learning per ottenere un trattamento ottimale analizzando numerosissime (spesso non ottimali) decisioni cliniche.
- Un modello puo' predire la diuresi post-rianimazione e identificare pazienti che saranno oligurici.
- I modelli AI "allenati" utilizzando parametri di ecocardiografia dimostrano capacita' di valutare lo stato volemico e la responsivita' ai fluidi nei pazienti critici con sensibilita' sovrapponibili al test del passive leg raising



XIII congresso nazionale

The National Heart, Lung, and Blood Institute Prevention and Early Treatment of Acute Lung Injury Clinical Trials Network. Early restrictive or liberal fluid management for sepsis-induced hypotension. N Engl J Med 2023;388:499–510.

Komorowski M, Celi LA, Badawi O, Gordon AC, Faisal AA. The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. Nat Med 2018;24:1716–20.

Lin PC, Huang HC, Komorowski M, Lin WK, Chang CM, Chen KT, et al. A machine learning approach for predicting urine output after fluid administration. Comput Methods Progr Biomed 2019;177:155–9.

Bataille B, Selle Jde, Moussot P-E, Marty P, Silva S, Cocquet P. Machine learning methods to improve bedside fluid responsiveness prediction in severe sepsis or septic shock: an observational study. Br J Anaesth 2021;126:826–34.

5) Stratificazione prognostica nella sepsi

- La valutazione prognostica e' fondamentale nella sepsi e l'AI ha mostrato dei vantaggi rispetto agli scores tradizionali.
- Un trial randomizzato controllato ha indicato che l'applicazione di modelli di AI **riduceva la mortalita' intraospedaliera dal 21.30 % all'8.96 %** confrontato con il gruppo di controllo insieme ad una **riduzione della degenza media in ospedale da 13.0 a 10.3 giorni**
- Uno studio multicentrico retrospettivo che includeva pazienti non di terapia intensiva (ICU) indicava che negli ospedali in cui erano utilizzati modelli di AI c'era una **minore mortalita', minore tasso di ricovero in ICU**, ridotti tempi di degenza confrontato con ospedali tradizionali

Shimabukuro DW, Barton CW, Feldman MD, Mataraso SJ, Das R. Effect of a machine learning-based severe sepsis prediction algorithm on patient survival and hospital length of stay: a randomised clinical trial. *BMJ Open Respir Res* 2017;4:e000234

Burdick H, Pino E, Gabel-Comeau D, McCoy A, Gu C, Roberts J, et al. Effect of a sepsis prediction algorithm on patient mortality, length of stay and readmission: a prospective multicentre clinical outcomes evaluation of real-world patient data from US hospitals. *BMJ Health Care Inform* 2020;27:e100109.

Escobar GJ, Liu VX, Schuler A, Lawson B, Greene JD, Kipnis P. Automated identification of adults at risk for in-hospital clinical deterioration. *N Engl J Med* 2020;383:1951–60

Yin J, Ngiam KY, Teo HH. Role of artificial intelligence applications in real-life clinical practice: systematic review. *J Med Internet Res* 2021;23: e25759.

Effect of a sepsis prediction algorithm on patient mortality, length of stay and readmission: a prospective multicentre clinical outcomes evaluation of real-world patient data from US hospitals

Table 3 Sepsis-related patient outcomes table—analysis of in-hospital mortality, hospital length of stay and 30-day readmissions, in the baseline and MLA periods for sepsis-related patient

	Baseline period	MLA period	Reduction
In-hospital mortality	3.86%	2.34%	39.50%
Length of stay	4.83 days	3.27 days	32.27%
30-day readmission	36.4%	28.12%	22.74%

There were 12 793 patients in the baseline period, of whom, 3592 were included for analysis and 62 354 patients in the MLA period, of whom, 14 166 patients were included for analysis
MLA, machine learning-based algorithm.

Uno altro studio multicentrico basato su dati in real-world suggerisce che l'applicazione di modelli di AI a pazienti in reparti di degenza e ED correla con una **riduzione della mortalità' intraospedaliera (del 39,5%), dei tempi di degenza (del 32,3%) e dei reingressi a 30 gg (del 22,7%)**

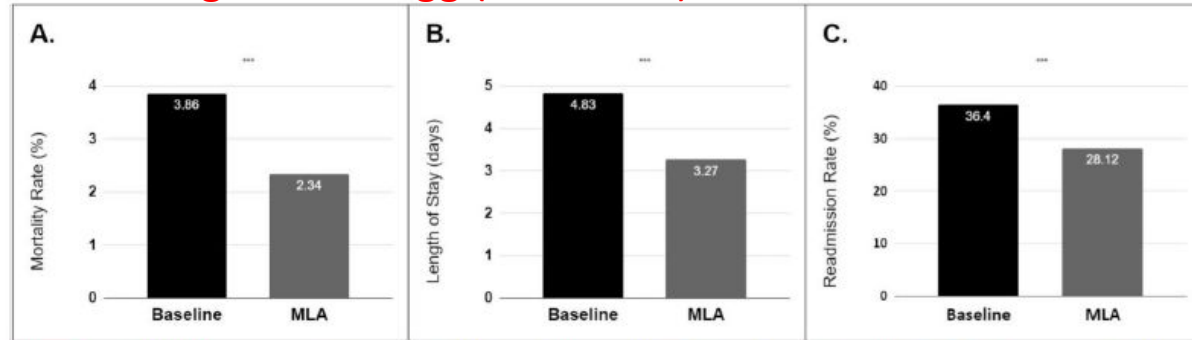


Figure 1 Patient outcomes—differences in (A) in-hospital mortality, (B) hospital length of stay and (C) 30-day readmissions in the baseline period and the MLA period for sepsis-related patients. Use of the MLA was associated with a 39.5% reduction of in-hospital mortality ($p < 0.001$), a 32.3% reduction in length of stay ($p < 0.001$) and a 22.7% reduction in 30-day readmissions ($p < 0.001$). MLA, machine learning-based algorithm.

Burckid H et al. Effect of a sepsis prediction algorithm on patient mortality, length of stay and readmission: a prospective multicentre clinical outcomes evaluation of real-world patient data from US hospitals. *BMJ Health Care Inform* 2020;27:e100109.

ARTICLE OPEN



Impact of a deep learning sepsis prediction model on quality of care and survival

L'implementazione in 2 ED di un modello di deep learning in real time per predire la sepsi era associato ad un **incremento del 5.0% della compliance al bundle della sepsi** e ad una **riduzione dell' 1.9% della mortalita' intraospedaliera..**

L'utilizzo di questi modelli era associato ad un miglioramento di outcome intermedi come una **riduzione delle insufficienze d'organo a 72 ore** che potrebbero spiegare gli effetti evidenziati. Inoltre dove gli infermieri riportano l'alert del sospetto di sepsi, gli antibiotici vengono somministrati prima.

Esempi di modelli di ML per previsione della sepsi

Emergency (1)

Could it be Sepsis?

Intervention Required (2nd 8 hours)

- Blood cultures before antibiotics
- Antibiotics
- Lactate level
- 30 ml/kg IV fluid bolus if hypotension or lactate > 4.0

Intervention Required (2nd 8 hours)

- Repeat Lactate if initial Lactate > 2
- Vasopressors if hypotension after fluids
- Repeat Physio of Lungs

Sepsis
SIRS + infection
Severe Sepsis + Organ Dysfunction

Severe Sepsis
Severe Sepsis + Organ Dysfunction

Septic Shock
Severe Sepsis + Organ Dysfunction

This patient has a Sepsis Risk Score: 90% chance of developing severe sepsis in the next 4 hours.

Consider discussing risk of sepsis with the primary physician or activating Code Sepsis

Top reasons in the past 6 hours
Sepsis Top Causes: Temperature, Heart Rate

[Secure Chat the Physician and Dr. Gabe Wardi](#)

The following actions have been applied:

✓ Sent: A summary of this advisory has been sent as a push notification

Acknowledge Reason

© 2023 Epic Systems Corporation

Algoritmi in uso

Sepsis Watch e SPOT

SPOT (Sepsis Prediction and Optimization of Therapy) (HCA Health Care)

- I computer, allenati gestendo milioni di dati di pazienti che hanno o non hanno sviluppato la sepsi, monitorizzano ogni secondo i dati provenienti dai pazienti ospedalizzati.
- Quando viene evidenziato un pattern di dati consistenti con il rischio di sepsi il computer dà un alert ai sanitari che attivano in “codice sepsi”. L’infermiere risponde valuta il paziente e se la sepsi non è esclusa inizia il trattamento immediatamente.

SPOT

How HCA Healthcare's algorithm-driven technology is spotting sepsis and saving lives

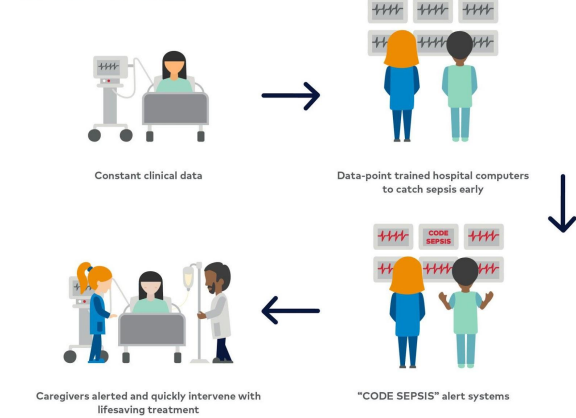
Sepsis Prediction and Optimization of Therapy = SPOT

Informed by data from millions of hospitalizations, SPOT was created by HCA Healthcare clinical and IT experts.



HCA
Healthcare

How SPOT Works:



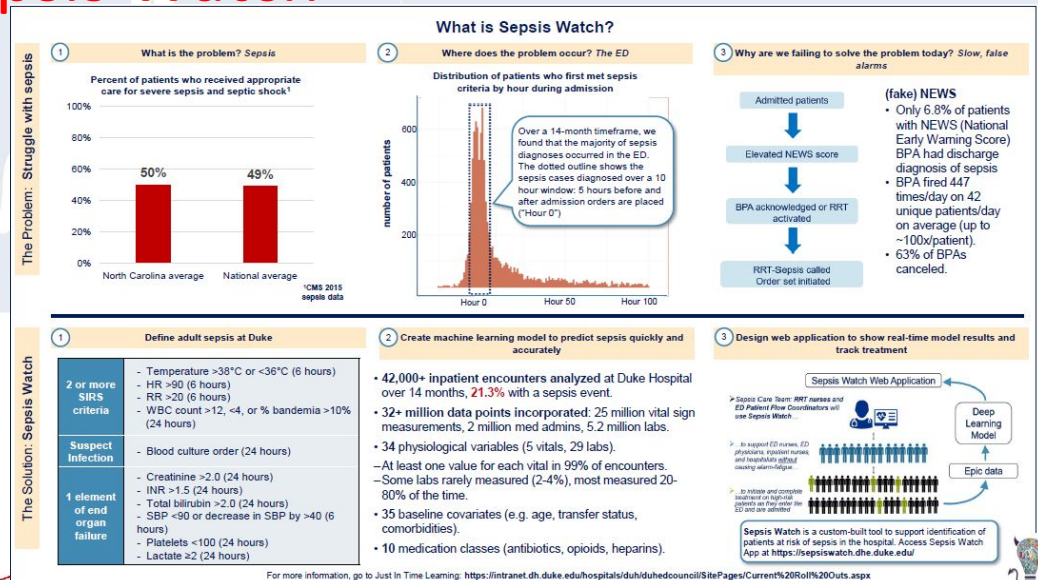
XXI Congresso Nazionale

L'utilizzo di questo programma ha permesso di identificare e di trattare i pazienti con un anticipo di 6 ore rispetto alla pratica comune. C'è stata una riduzione della mortalità rispetto all'anno precedente all'introduzione di SPOT del 9.9% per la sepsi severa e del 5.1% per lo shock settico

Algoritmi in uso

Sepsis Watch

Alla Duke University e' stato sviluppato ed implementato un modello di deep learning che da alert precoce per i pazienti a rischio di sepsi presenti in ED. Da quando e' in uso il Sepsis Watch™ dal novembre 2018 la Duke ha raddoppiato la compliance al bundle delle 3 ore della SSC



Sepsis Watch analizza dati come i parametri vitali, risultati degli esami, comorbidita', dati demografici, dati storici del paziente ogni 15'. Analizza 86 variabili differenti, li campiona e li combina e evidenzia relazioni tra le variabili che potrebbero essere segnali di sepsi iniziale. Se un paziente ha i criteri di esordio della sepsi il sistema lo segnala al team di emergenza intraospedaliera.

- Concetti generali
- Applicazioni nei pazienti con sepsi in Medicina d'urgenza
- Aspetti controversi



Limiti tecnici

Anche se la ML offre un elevato potenziale nella ricerca clinica e' assolutamente necessario conoscerne i limiti.

- Primo, **l'aspetto della causalita'** e' una sfida perche' i dati correlati possono essere erroneamente costruiti come se ci fosse un nesso di causalita'.
- Secondo, **la riproducibilita' degli studi** sulla ML e' un aspetto critico vista la complessita' della natura degli algoritmi e la dipendenza dalla qualita' dei dati.
- Infine la **mancaanza di trasparenza nella metodologia e nell'uso dei database** puo' non favorire la replicabilita' e la comprensione degli studi.

Queste sfide richiedono uno sforzo congiunto di comunicazione scientifica e trasparenza metodologica per mantenere l'integrita' e l'utilita' della ricerca in ML.



HHS Public Access

Author manuscript

Annu Rev Biomed Data Sci. Author manuscript; available in PMC 2022 July 01.

Published in final edited form as:

Annu Rev Biomed Data Sci. 2021 July ; 4: 123–144. doi:10.1146/annurev-biodatasci-092820-114757.

Ethical Machine Learning in Healthcare

Aspetti etici

- Via via che I modelli di machine learning (ML) proliferano in molti aspetti delle nostre vite, c'è grande preoccupazione che possano recare danno. Una delle preoccupazioni in ambito medico è legata ad aspetti di natura etica, per il potenziale di questi strumenti di **peggiore disparità di diritto alla salute** già esistenti
- Un recente lavoro ha dimostrato che modelli predittivi hanno una bassa performance sulle donne , minoranze etniche e razziali e su quelli con assicurazioni pubbliche



Topol EJ. 2019. High-performance medicine: the convergence of human and artificial intelligence. *Nat. Med*25:44–56 [PubMed: 30617339]

Ferryman K, Winn RA. 2018. Artificial intelligence can entrench disparities—Here's what we must do. *The Cancer Letter*, 11. 16. https://cancerletter.com/articles/20181116_1/

Wiens J, Saria S, Sendak M, Ghassemi M, Liu VX, et al. 2019. Do no harm: a roadmap for responsible machine learning for health care. *Nat. Med*25:1337–40 [PubMed: 31427808]

Ghassemi M, Naumann T, Schulam P, Beam AL, Chen IY, Ranganath R. 2020. A review of challenges and opportunities in machine learning for health. *AMIA Summits Transl. Sci. Proc*2020:191–200 [PubMed: 32477638]

Ghassemi M, Naumann T, Schulam P, Beam AL, Chen IY, Ranganath R. 2019. Practical guidance on artificial intelligence for health-care data. *Lancet Digital Health*1:e157–59 [PubMed: 33323184]

Chen IY, Szolovits P, Ghassemi M. 2019. Can AI help reduce disparities in general medical and mental health care? *AMA J. Ethics*21:167–79



HHS Public Access

Author manuscript

Annu Rev Biomed Data Sci. Author manuscript; available in PMC 2022 July 01.

Published in final edited form as:

Annu Rev Biomed Data Sci. 2021 July ; 4: 123–144. doi:10.1146/annurev-biodatasci-092820-114757.

Ethical Machine Learning in Healthcare

- Una mole crescente di letteratura sta affrontando le implicazioni sociali della ML. Alcuni di questi lavori chiamati “critical data studies”, valutano il punto di vista clinico mentre altri lavori valutano quello tecnico e di computer science

Boyd D, Crawford K. 2012. Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon. *Inform. Commun. Soc*15:662–79

Dalton CM, Taylor L, Thatcher J. 2016. Critical data studies: a dialog on data and space. *Big Data Soc*. 3(1). 10.1177/2053951716648346

Zliobaite I2015. A survey on measuring indirect discrimination in machine learningarXiv:1511.00148 [cs.CY]

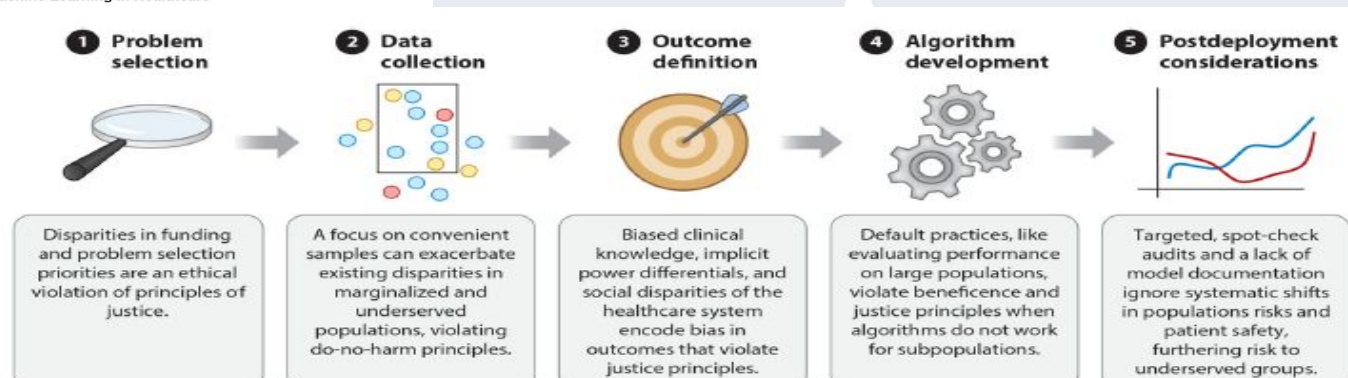
Barocas S, Hardt M, Narayanan A. 2018. Fairness and machine learning. Online Book, fairmlbook.org. <http://www.fairmlbook.org>

Corbett-Davies S, Goel S. 2018. The measure and mismeasure of fairness: a critical review of fair machine learning. arXiv:1808.00023 [cs.CY]





Ethical Machine Learning in Healthcare



Analisi del processo:

- Primo: **selezione del problema** (I fondi per la ML possono portare ad ingiustizie nella scelta).
- Secondo: **processi di campionamento** nelle ricerche sponsorizzate possono amplificare l'ineguaglianza
- bias legati alla **scelta dell'outcome** o alla **costruzione dell'algoritmo** (ci possono essere molti fattori che possono influenzare la performance del modello che puo' non essere valido in sottopopolazioni)
- Audit post implementazione non adeguati

Explainable artificial intelligence in emergency medicine: an overview

eISSN: 2383-4625

Trasparenza e “spiegabilità”

- Spesso I clinici che non hanno esperienza di AI la percepiscono come una “scatola nera”



Explainable artificial intelligence in emergency medicine: an overview

eISSN: 2383-4625

- Nonostante si pensi che I medici d'urgenza diventeranno grandi utilizzatori di AI e ML nel prossimo futuro, gli scettici sono dubbiosi che I clinici si fidino e si basino su modelli di AI e ML
- L'obiettivo dell'AI “spiegabile” e' di insegnare il funzionamento dei sistemi di AI agli utilizzatori
- I concetti chiave sono quelli della trasparenza e della “spiegabilita”

XIII congresso nazionale

Minh D, Wang HX, Li YF, Nguyen TN. Explainable artificial intelligence: a comprehensive review. *Artif Intell Rev* 2022;55:3503–68.

Arrieta AB, Diaz-Rodriguez N, Del Ser J, et al. Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf Fusion* 2020;58:82–115.

. Doran D, Schulz S, Besold TR. What does explainable AI really mean? A new conceptualization of perspectives [Preprint]. Posted 2017 Oct 2. *arXiv:171000794*.

<https://doi.org/10.48550/arXiv.1710.00794>

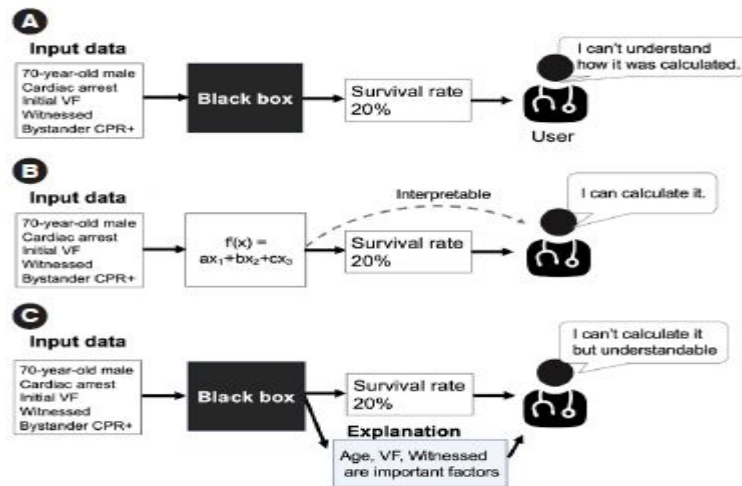
Explainable artificial intelligence in
emergency medicine: an overview

Fig. 1. Black box, transparency, and post hoc explainability. (A) Black-box model. (B) Transparent model. (C) Post hoc explainability. VF, ventricular fibrillation; CPR, cardiopulmonary resuscitation.

Trasparenza: i modelli di
AI sono spesso etichettati
con “scatola nera”
suggerendo una carenza
di trasparenza come
mostrato in Fig. 1A

- La “spiegabilita’” e’ un concetto differente
- Arrieta et al. suggerivano che , “Dato un certo pubblico la “spiegabilita’” si riferisce all’abilita’ di un modello di mostrare dettagli e di rendere chiaro o facile da comprendere il suo funzionamento interno.”



Arrieta AB, Diaz-Rodriguez N, Del Ser J, et al. Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Inf Fusion 2020;58:82-115.



GENOVA 30 MAG - 1 GIU 2024

- Ci sono 4 principali motivazioni perche' e' necessario che l'AI sia spigabile in ambito clinico
- **giustificazione:** giustificare gli output del modello di AI per favorire la fiducia e per supportare la clinical decision-making (I clinici devono spiegare la condizione clinica, il piano di trattamento e gli outcome previsti ai pazienti e familiari)
- **Controllo:** la "spiegabilita" aiuta gli utilizzatori a mantenere il controllo di una tecnologia complessa (una conoscenza piu' approfondita aumenta la consapevolezza delle vulnerabilita' e difetti e permette di identificare rapidamente e correggere errori in situazioni critiche)
- **Miglioramento** la "spiegabilita" di modelli di AI e' necessaria per favorire miglioramento continuo (se I modelli di AI possono mostrare il processo utilizzato per produrre specifici risultati l'informazione puo' essere utilizzata per ulteriori miglioramenti).
- **Scoperta di nuove idee:** puo' permettere lo sviluppo di nuove idee, ipotesi (se una spiegazione da un modello di AI mostra un contributo inatteso da parte di un certo fattore di rischio per la previsione di un outcome, si puo' sviluppare ~~una nuova ipotesi~~ relativamente a quel fattore di rischio e la sua associazione con l'outcome)

XIII congresso nazionale

Adadi A, Berrada M. Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). IEEE Access 2018;6:52138-60.



GENOVA 30 MAG - 1 GIU 2024

La “spiegabilità” dell’AI comprende 3 approcci principali:

- Il primo è il “**premodeling**” inteso come visualizzazione, riepilogo e trasformazione dei dati
- Il secondo approccio è **sviluppare un modello interpretabile con comprensibilità intrinseca** : I modelli possono avere vari livelli di interpretabilità e trasparenza: a livello dell’algoritmo di training (chiamato “algorithmic transparency”), al livello dei componenti (noto come “decomposability”), e al livello del modello stesso (o “simulatability”)
- L’ultimo metodo è chiamato **post-modeling explainability** .Aiuta a scomporre modelli AI complessi in modo che siano più facili da comprendere. Queste tecniche sono state create basandosi sulle capacità di comprensione umana

XIII congresso nazionale
simeu

Minh D, Wang HX, Li YF, Nguyen TN. Explainable artificial intelligence: a comprehensive review. Artif Intell Rev 2022;55:3503–68.

Arrieta AB, Diaz-Rodriguez N, Del Ser J, et al. Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. Inf Fusion 2020;58:82–115.

Doran D, Schulz S, Besold TR. What does explainable AI really mean? A new conceptualization of perspectives [Preprint]. Posted 2017 Oct 2. arXiv:171000794.

<https://doi.org/10.48550/arXiv.1710.00794>

Valutazione della qualità

- Le linee guida correnti sulla AI non danno un'idea della qualità della ricerca sulla AI, limitando la capacità di confronto di modelli che rispondono allo stesso quesito clinico.
- Questo studio ha valutato modelli di AI utilizzando lo strumento APPRAISE-AI, un metodo quantitativo per valutare la qualità degli studi di AI su 6 aspetti: rilevanza clinica, qualità dei dati, condotta metodologica, robustezza dei risultati, qualità dei report e riproducibilità.



Ogni aspetto dell' APPRAISE-AI e' stato mappato in uno dei seguenti domini: **rilevanza clinica, qualita' dei dati, condotta metodologica, robustezza dei risultati, qualita' del reporting, e riproducibilita'**

Gli scores devono essere sommati per determinare la qualita' complessiva dello studio (score torale di APPRAISE-AI)

Table 1. APPRAISE-AI Domains and Corresponding Items^a

Domain and items	Domain score
Clinical relevance	
1. Title	
2. Background	
3. Objective and problem	4
21. Implementation into clinical practice	
Data quality	
4. Source of data	
5. Eligibility criteria	
6. Ground truth	24
7. Data abstraction, cleaning, preparation	
Methodological conduct	
8. Data splitting	
9. Sample size calculation	20
10. Baseline	
Robustness of results	
15. Model evaluation	
16. Clinical utility assessment	
17. Bias assessment	20
18. Error analysis	
19. Model explanation	
Reporting quality	
13. Cohort characteristics	
20. Critical analysis	
22. Limitations	12
23. Disclosures	
Reproducibility	
11. Model and processing description	
12. Hyperparameter tuning	
14. Model specification	20
24. Transparency	
Overall score	100

^a Please refer to eTable 1 in Supplement 1 for a detailed breakdown of each item. The overall APPRAISE-AI score was graded as follows: very low quality, 0 to 19; low quality, 20 to 39; moderate quality, 40 to 59; high quality, 60 to 79; and very high quality, 80 to 100.

APPRAISE-AI ha dimostrato un'alta affidabilità ed ha correlato bene con molte misure di qualità degli studi. Questo strumento potrebbe essere utilizzato dagli sperimentatori, reviewers, editori e sponsors per valutare la qualità degli studi sulla AI da utilizzare per il supporto decisionale.



Conclusioni

- L'AI e la ML sono diventate parte comune nella ricerca e nella pratica sanitaria
- L'AI potrebbe aiutare nelle decisioni al triage, nella gestione e nella stratificazione prognostica dei pazienti con sepsi
- E' necessario che i ricercatori sviluppino modelli sempre piu' comprensibili da parte dei clinici per favorirne la diffusione.
- D'altro lato I clinici devono stabilire una forte alleanza con partner con abilita' particolari in computational science
- Sono necessari piu' studi "real world" che confrontano AI e "standard of care" o giudizio clinico.