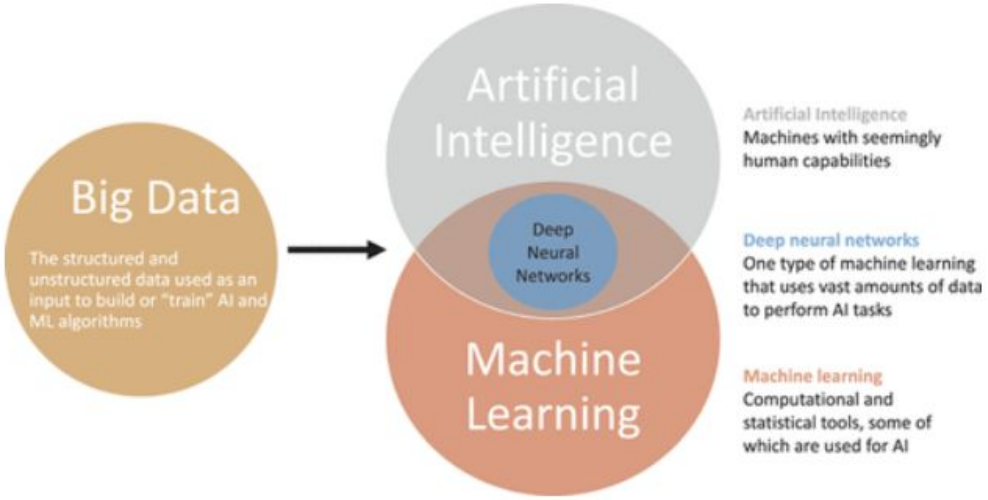
eduARCAset

Incontro su interpretazione del GRT, di stampo didattico.
Evoluzione del test di resistenza per HIV - Implicazioni cliniche e laboratoristiche

15 FEBBRAIO 2023

I-PRO

euresist network



ARCAset nasce come un "progetto" dell'edizione 2020 dell'ARCA Factory
Il corso-laboratorio – la Factory – si articola in tre fasi distinte e consequenziali:

- *Valutazione dei temi che necessitano di indagine nella gestione dell'infezione da HIV*
- *Disegno di progetti utili a risolvere le carenze identificate*
- *Interrogazione delle basi dati utili e analisi delle strategie di ricerca dei dati idonei allo svolgimento del progetto*

Il Tema del gruppo di lavoro era:
costituzione di Repository di casi clinici che possa essere di supporto alle decisioni diagnostiche e terapeutiche

il progetto è cresciuto: non solo Repository, ma formazione sul campo e utilizzo di tecniche di AI per dare indicazioni basate sui dati

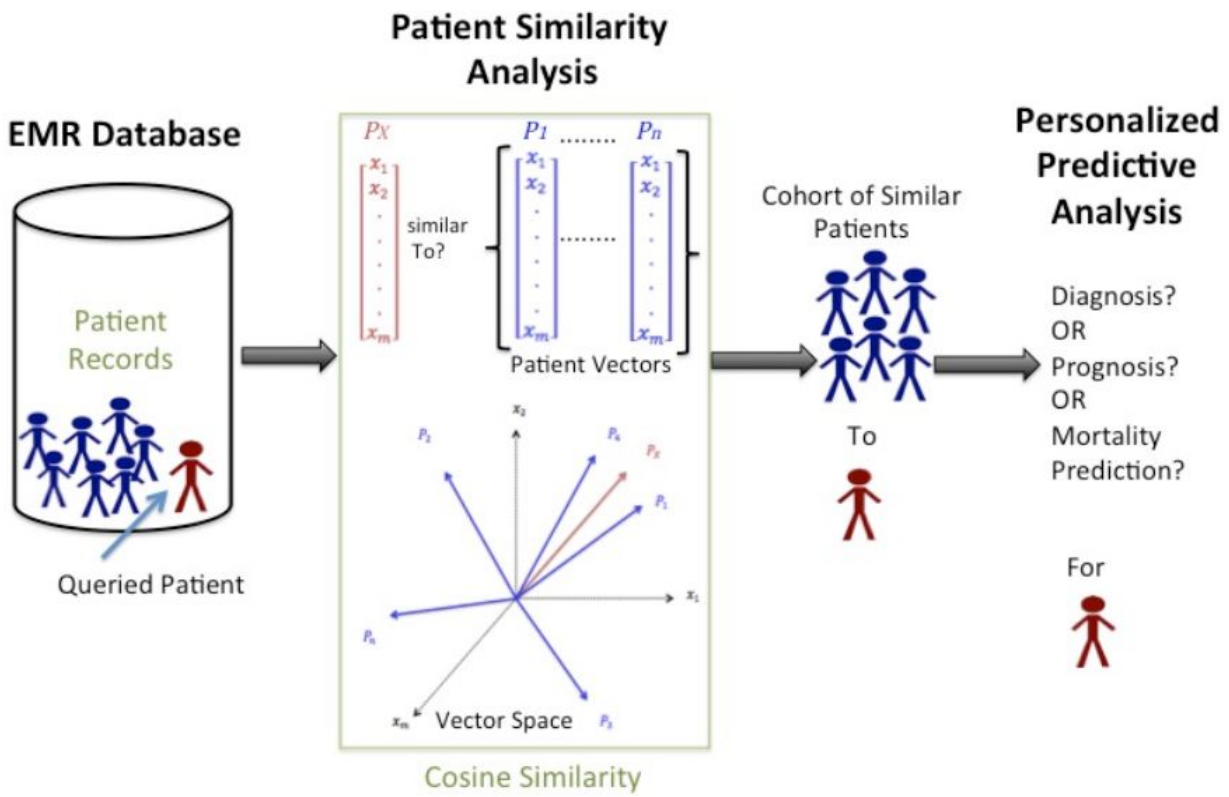
eduARCAset
Incontro su interpretazione del GRT, di stampo didattico.
Evoluzione del test di resistenza per HIV - Implicazioni cliniche e laboratoristiche

Cercare casi clinici in un database è un problema comune e affrontato da diversi punti di vista:

- *Cercare dati predittivi di patologia definita come caso clinico standard*
- *Cercare dati predittivi di diagnosi efficace*
- *Cercare in database non clinici elementi predittivi di patologia*

…. Il nostro caso può meglio inserirsi nella definizione:
Analisi della somiglianza dei pazienti all'interno di un database,
per un sistema personalizzato di supporto alle decisioni mediche

Ci sono molti approcci possibili



Efficient In-Database Patient Similarity Analysis for Personalized Medical Decision Support Systems

Araek Tashkandi[a,1,*], Ingmar Wiese[a], Lena Wiese[a]

[a] Georg-August-University Goettingen, Institute of Computer Science, Goldschmidtstr. 7, 37077 Goettingen, Germany.

L'analisi della somiglianza del paziente è una precondizione per applicare la tecnologia di apprendimento automatico ai dati clinici. In questo senso, l'analisi della somiglianza tra i pazienti sfrutta la ricchezza di informazioni delle cartelle cliniche elettroniche (EMR) per supportare il processo decisionale medico. Un calcolo di somiglianza a coppie può essere utilizzato come base

La letteratura si popola rapidamente di molti articoli centrati Sulla ricerca di "casi simili"

Study	Patient Similarity for	Patient Similarity Metric	Evaluation Metrics	Implementation Tool
Sun et al. [19]	Prognosis prediction of occurrence of Acute Hypotensive Episode (AHE)	Locally Supervised Metric Learning (LSML), Mahalanobis distance	Accuracy of: Classification, Retrieval, and Prognosis	Developed system
Wang et al. [21]	Disease prognosis	Euclidean distance	Classification performance: precision, Recall, Accuracy, and F-measure	Developed SimProX system
Morid et al. [20]	Mortality prediction	Classification with K-Nearest Neighbor algorithm	Classification performance: precision, Recall, and F-measure	Developed PPMF Framework
Lee et al. [4]	Mortality prediction	Cosine Similarity metric	Accuracy: Area under ROC curve; area under precision-recall curve	R
Hoogendoorn et al. [15]	Mortality prediction	K-Nearest Neighbor, Euclidean distance	AUC for: -Influence of the number of patients upon the predictions accuracy and the computation time. -Earliest Prediction Time	Python
Panahiazar et al. [22]	Medication plan	Supervised and unsupervised clustering approaches with Mahalanobis distance	Sensitivity, specificity, F-measure, and accuracy	Developed framework
Gottlieb et al. [5]	Predict the eventual discharge diagnoses	Jaccard, Euclidean distance	AUC and F-measure	MATLAB

Application of Patient Similarity in Smart Health: A Case Study in Medical Education

Kalkidan Fekadu Eteffa, Samuel Ansong, Chao Li, Ming Sheng, Yong Zhang & Chunxiao Xing

Conference paper | First Online: 16 September 2019

1776 Accesses | 1 Citations

Part of the Lecture Notes in Computer Science book series (LNISA,volume 11817)

Abstract

Patient similarity relies on computations that synthesize EHRs (Electronic Health Records) to give personalized predictions, which inform diagnoses and treatments. Given the complexities in pre-processing EHRs, representing patient data and utilizing the most suitable similarity metrics and evaluation methods, patient similarity computations are far from the era of regular use in hospitals. This paper aims to both support further patient similarity research and to inform the importance of its application in medical education. It accomplishes this by examining relevant literature that offer techniques to tackle the computational challenges and by presenting their various applications in the healthcare industry.

Date le complessità nella pre-elaborazione delle cartelle cliniche elettroniche, nella rappresentazione dei dati dei pazienti e nell'utilizzo delle metriche di somiglianza e dei metodi di valutazione più adatti, i calcoli della somiglianza dei pazienti sono lontani dall'era dell'uso regolare negli ospedali

International Journal of Medical Informatics
Volume 135, March 2020, 104073

A patient-similarity-based model for diagnostic prediction

Zheng Jia [a], Xian Zeng [a], Huilong Duan [a], Xudong Lu [a], Haomin Li [b]

Abstract

Objective

To simulate the clinical reasoning of doctors, retrieve analogous patients of an index patient automatically and predict diagnoses by the similar/dissimilar patients.

Methods

We proposed a novel patient-similarity-based framework for diagnostic prediction, which is inspired by the structure-mapping theory about analogy reasoning in psychology. Patient similarity is defined as the similarity between two patients' diagnoses sets rather than a dichotomous (absence/presence of just one disease). The multilabel classification problem is converted to a single-value regression problem by integrating the pairwise patients' clinical features into a vector and taking the vector as the input and the patient similarity as the output. In contrast to the common k-NN method which only considering the nearest neighbors, we not only utilize similar patients (positive analogy) to generate diagnostic hypotheses, but also utilize dissimilar patients (negative analogy) are used to reject diagnostic hypotheses.

la somiglianza del paziente è definita come
la somiglianza tra i set di diagnosi/dati clinici di due pazienti
piuttosto che una dicotomia
(assenza/presenza di una sola malattia).
Il problema di classificazione multietichetta
viene convertito in un problema di regressione
a valore singolo integrando le caratteristiche cliniche
dei pazienti a coppie in un vettore
e prendendo il vettore come input
e la somiglianza del paziente come output.

IL NOSTRO PUNTO DI PARTENZA E' IL dbARCA

si tratta di un database clinico con molte informazioni

DATI PRESENTI:

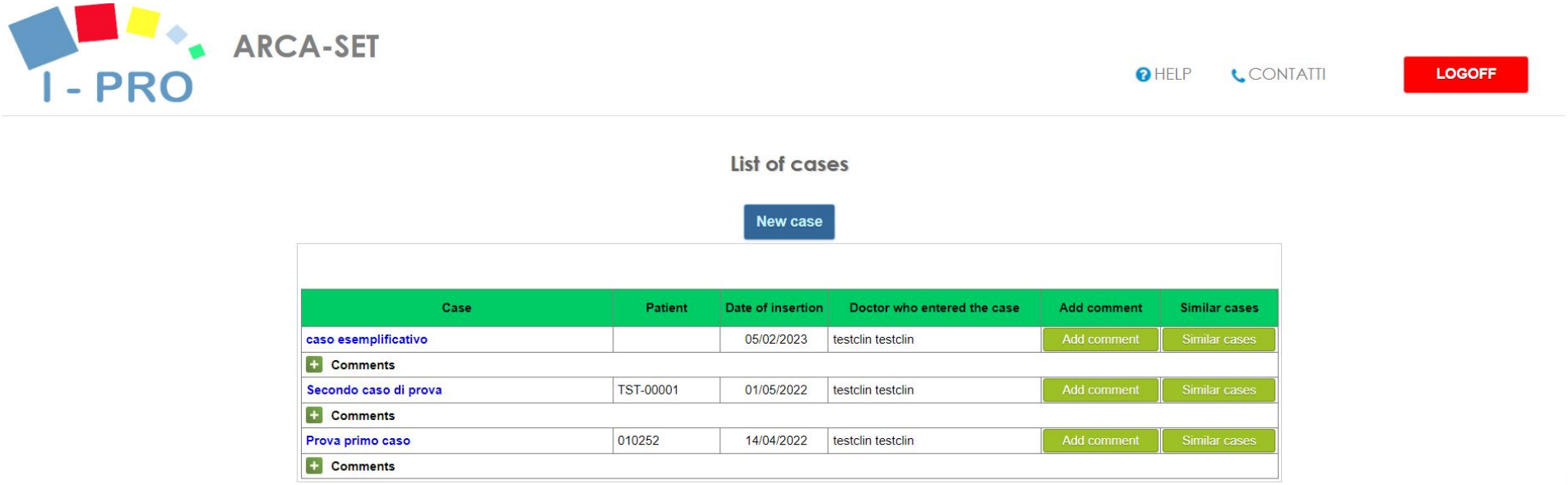
>47.000 pazienti
>58.000 genotipi
>89.000 conte cd4
>71.000 viremie
>133.000 terapie (intese come combinazioni di farmaci)

DETTAGLIO DEI DATI:
(tables di dati del database)

- tbbatterimiceti
- tbbiopsia
- tbcd4
- tbcns
- tbcorecettore
- tbcredits
- tbcreditsold
- tbdexa
- tbdiario
- tbecografia
- tbelastometria
- tbencefalepatica
- tbesamelabor
- tbesamestrument
- tbeventiaids
- tbfiltratoglom
- tbgenetica
- tbgenotipo
- tbgenotipoautor
- tbgenotipoold
- tbgenotipopubl
- tbgenotiporapido
- tbgravidanze
- tbhla
- tbhlaold
- tbipersens
- tblaboratoriautor
- tbmessages
- tbmessagespubl
- tbmutresgenotip
- tbmutresgenotippubl
- tbmutresrtgenotip
- tbnaive
- tbnews
- tbnoneventoaids
- tbnotes
- tbnotespubl
- tboptions
- tbopzioni
- tbpatologie
- tbpazienti
- tbpazientiautor
- tbprecedente
- tbpropostoutente
- tbpubblicazioni
- tbpubblicazioniold
- tbreparti
- tbrepartiautor
- tbrichieste
- tbrichiesteold
- tbrm
- tbscore
- tbscorepubl
- tbsequenza
- tbsequenzapubl
- tbsubtypes
- tbsubtypespubl
- tbterapiaaltre
- tbtrapianto
- tbtratfarm
- tbtratfarmdett
- tbunit
- tbunitcasi
- tbutenti
- tbvaccinazione
- tbviremie
- tbvisitaspecial
- tbwarning
- tbwarningpubl

l'utente avrà una lista di casi clinici

- **o inseriti ex novo dagli utenti**
- **o costruiti dai dati presenti nel dbARCA**



ARCA-SET

HELP CONTATTI

Case registration

Title of case:	Prova primo caso
Patient:	010252 010252
Case:	Prova primo casoProva primo casoProva primo casoProva primo casoProva primo casoProva primo casoProva primo casoProva primo caso
Date of insertion:	14/04/2022
Doctor who entered the case:	testclin testclin
Department:	Centro TEST
Gender:	(•) M ( ) F ( ) T (Transgender)
Year of birth:	1951
Weight:	75 Kg
Height:	175 cm
Metabolic disease:	(•) Yes ( ) No
If **Yes**, specify:	[ ] Diabetes / glucose intolerance [ ] Dyslipidemia [x] Obesity [x] Lipodystrophy [ ] Other, specify
Psychiatric disorders:	( ) Yes (•) No
Sleep disorders:	(•) Yes ( ) No
HBV:	Positive (chronic carrier)
HCV:	Positive (chronic carrier)

Difficulty in swallowing:	( ) Yes ( ) No
HLAB57:	
Last creatinine value:	
Previous regimes and reason for suspension:	Drug / Start date / End date / Reason suspension APV / 01/01/2012 / 30/01/2012 / HIV intensificazione RAL+TPV / 18/05/2013 / 30/05/2015 / HIV tossicità FTC+RPV+TDF / 11/01/2016 /
Antiretroviral drugs that cannot be used:	Add drugs

Save Back

La descrizione del singolo caso clinico potrà prevedere informazioni aggiuntive rispetto a quelle presenti nel dbARCA

In particolare saranno generati i TAG di classificazione

Similar cases

Set tags:

Terapia

- Terapia — Select therapy
- Farmaci antiretrovirali che non possono essere utilizzati — Add drugs

Condizioni/Parametri clinici

- HBV
- HCV
- Scarsa compliance
- <200 CD4
- HIVRNA non rilevato
- HIVRNA <100.000
- Interazione farmacologica
- HLAB57
- Tosssicità/intolleranza
- Semplificazione — Min: Max:
- Nadir CD4 — Min: Max:
- Zenith RNA

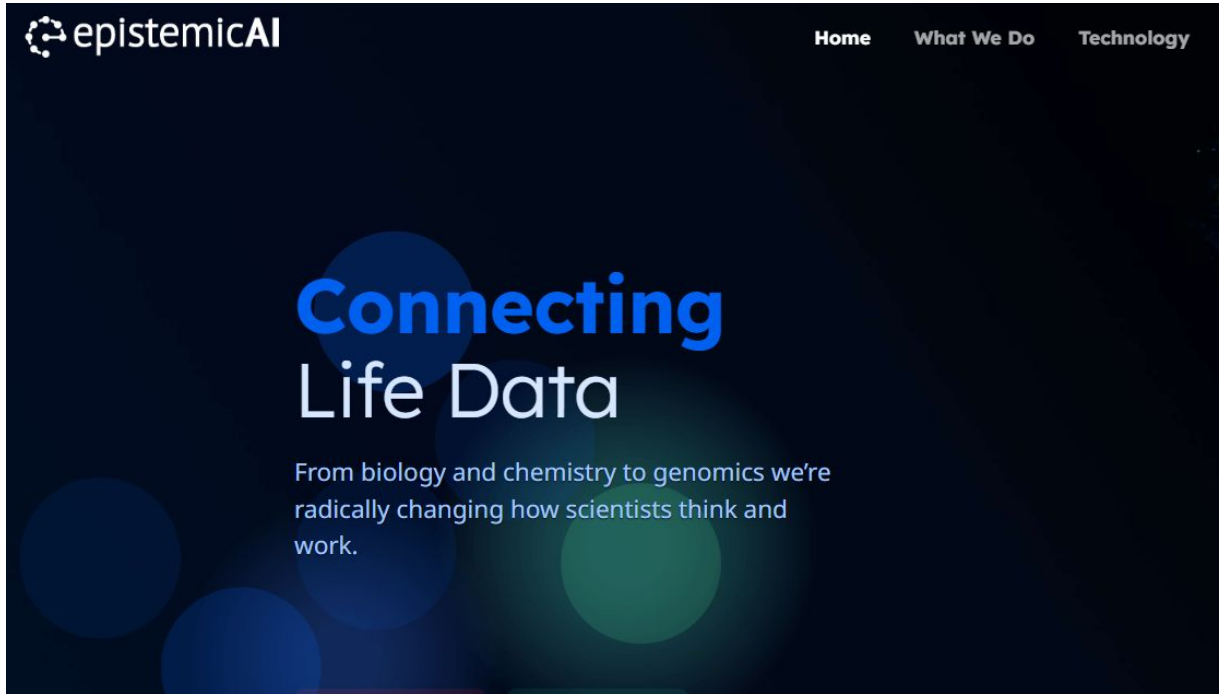
Mutazioni

- Protease Resistence — Select mutations
- Reverse Transcriptase Resistence — Select mutations
- Integrase Resistence — Select mutations
- Multi resistente — Select mutations

Condizioni patologiche/Anamnesi

- Obesità
- Patologia cardiovascolare
- Patologia polmonare
- Patologia reumatica/autoimmune
- Patologia renale

La ricerca dei casi "simili" potrà essere filtrati con ulteriori specifiche



Il sistema AI richiede:

grande potenza di calcolo
disponibilità di molto tempo/macchina

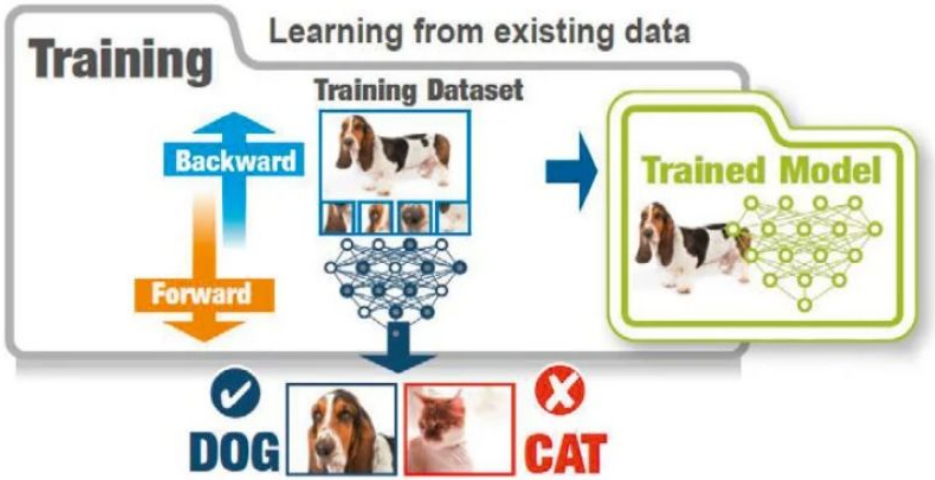
algoritmi e software che utilizzano tecniche di AI

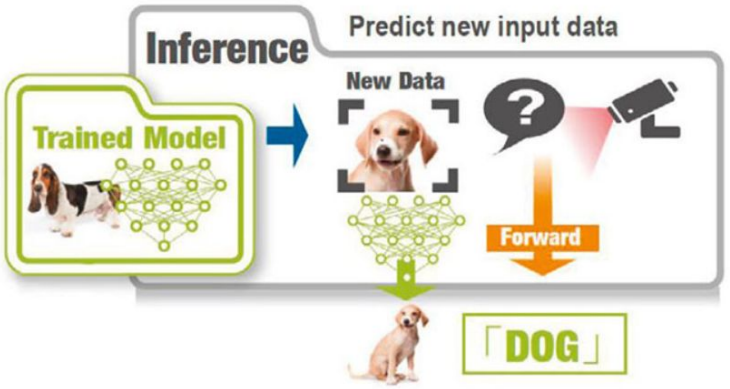
machine learning, reti neurali, etc…

L'utilizzo di metodologie statistiche tradizionali renderebbe la ricerca di casi "simili" più difficile e lunga

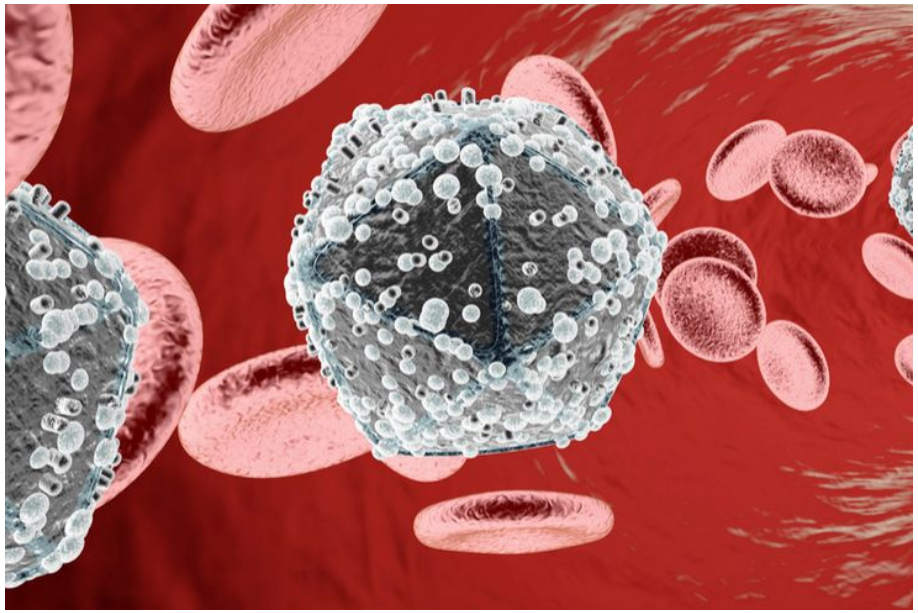
di quanto non possano fare le tecniche di AI

EURESIST sta valutando per ARCAset diversi possibili partner per lo sviluppo del motore di ricerca di casi "simili"



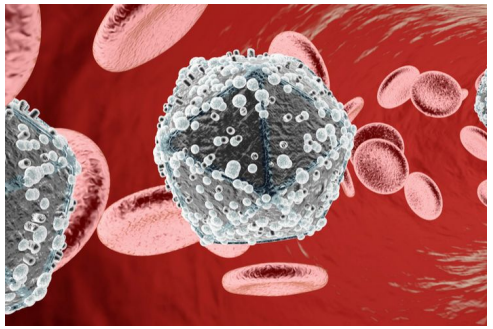


Ogni sistema basato su tecnica di IA
deve essere “addestrato”
partendo da dati reali
indicando regole e meccanismi di selezione
e ottenendo quindi un sistema che,
utilizzando le tecniche di IA che sono ormai oggetto di
continua ricerca e innovazione,
sia capace di dare risposte che vadano oltre
quanto esistente nei dati reali di partenza.



Gli aspetti virologici del caso clinico target
sono essenziali per la definizione di “similitudine”
e sono anche particolarmente complessi da definire
poiché il livello di approfondimento
può essere “grande a piacere”

l’interpretazione dei test
e la descrizione degli aspetti virologici
sono quindi un elemento del “linguaggio”
della comunicazione tra i ricercatori e i clinici
al tempo stesso critico e cruciale



In conclusione

Avevamo l'obiettivo di provare a sviluppare un sistema capace, dato un caso clinico di partenza, di generare un insieme limitato e "utile" di casi clinici simili per supportare un "consulto informato".

Abbiamo sviluppato una prima versione di sistema per la descrizione di un caso clinico (eventualmente a partire da dati già presenti in ARCA) e per l'avvio della ricerca dei casi clinici, con descrizioni anche testuali di cosa si sta cercando

Entro il 2023 potremo avere una prima versione del sistema di ricerca

….. il tutto nel rispetto delle norme su sicurezza e riservatezza dei dati…..

I gruppi di ricerca su AI (Google, Microsoft, Amazon, IBM, ……. e molti altri….) considerano il lavoro sui dati clinici come uno dei "bersagli" principali della loro attività di sviluppo

eduARCAset

Incontro su interpretazione del GRT, di stampo didattico.
Evoluzione del test di resistenza per HIV - Implicazioni cliniche e laboratoristiche