Available online at **www.joac.info**

ISSN: 2278-1862

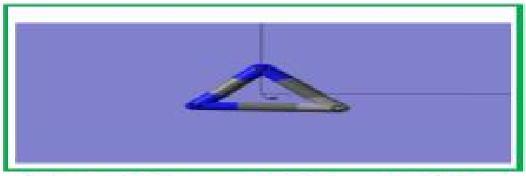
Journal of Applicable Chemistry

2023, 12 (4): 390-443

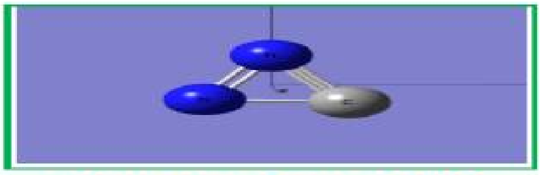
(International Peer Reviewed Journal)

New Chemistry News

$^{-}N=C=N^{-}$



New News of Chem (NNC)



ChemNewsNew (CNN)

CNN-53--Fit *(Figure Image TableScript...)* Bases (Bfit) 2022-2023 Part 1.xA I.Medicine (xAIM)

Information Source	**sciencedirect.com;**	
S. Narasinga Rao M D **Associate Professor,** **Dept. of General Medicine,** **Government medical college,** **government general hospital,** **Srikakulam,** AP, **India**	**K. Somasekhara Rao, Ph D** **Dept. of Chemistry,** **Acharya Nagarjuna Univ.,** **Dr. M.R.Appa Rao Campus,** **Nuzvid-521 201, India**	**R. Sambasiva Rao, Ph D** **Dept. of Chemistry,** **Andhra University,** **Visakhapatnam 530 003,** **India**
snrnaveen007@gmail.com (+91 9848136704)	sr_kaza1947@yahoo.com (+91 98 48 94 26 18)	rsr.chem@gmail.com (+91 99 85 86 01 82)

Conspectus: Passive databases of last century given way to computer readable ones in the scientific world. Spectral bases popular in chemistry/physics added a new dimension in interpretation of processes, properties and responses inmaterials/energy domains.The multi-dimensional-images brought renaissance in medical diagnosis. The knowledge bases emerged from distillation of information and intermediate hypotheses include logical/literature/numerical data. Typical case studies dealing with xAI in medical diagnosis during 2022 and 2023 are briefly described.

***Keywords*:**; Post-hoc explanations; Heatmaps; Saliency maps; xAI-Probes; Shapley; LIME; CAM; Grad-CAM; Integrated gradients; Class Activation Maps; tSNE plot; eXplainable/ interpretable/

Layout		
Diagnosis	Cancer	
	Heart diseases ECG Analysis	
	ASD	
	COVID-19	
Health	Drugs	K(nowledge)Lab rsr.chem1979
	Toxicology	
	Health-care	
xAI	Framework-	
	Segmentation	
	Explainability Interpretability	
	Pixel level	

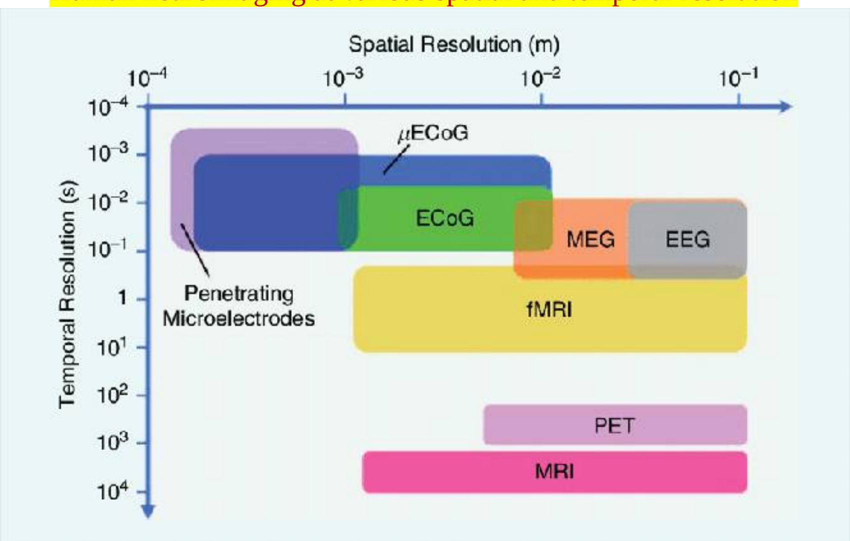
Diagnosis of Diseases (DoD)

Resolution of probes (Medical instruments)

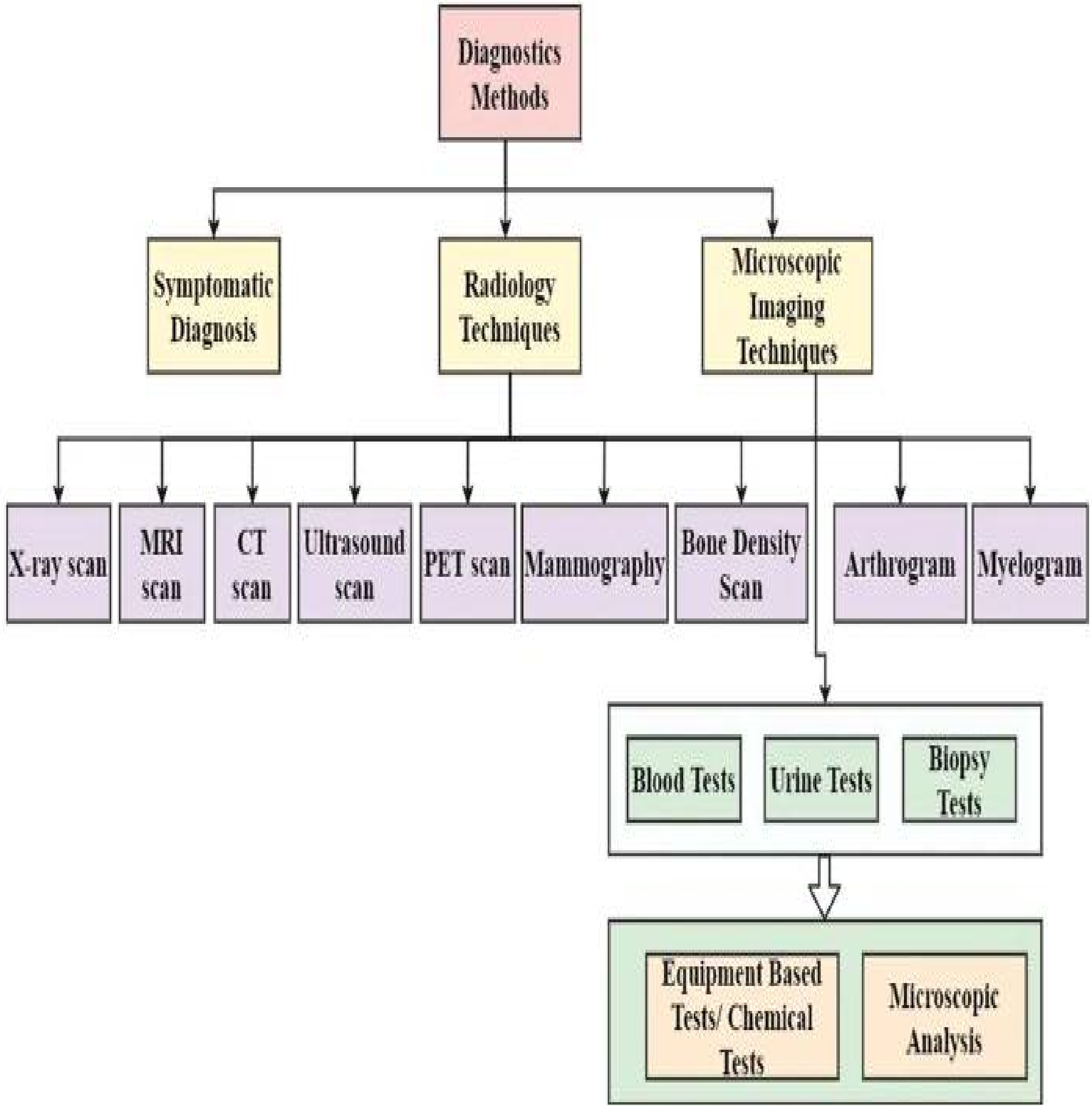
Spatial	Temporal

xAI.Med.		2022-10

Human neuroimaging at various spatial and temporal resolution



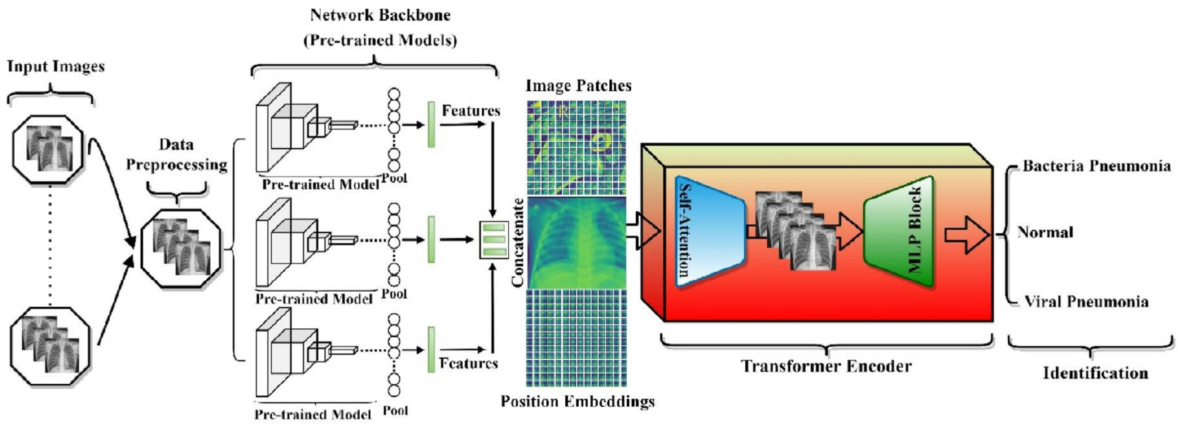
xAI.Med.	Diagnosis methods	2022



Lungs

xAI.Med. | **Case Study** | 2022-07

Pneumonia identification from chest X-ray images

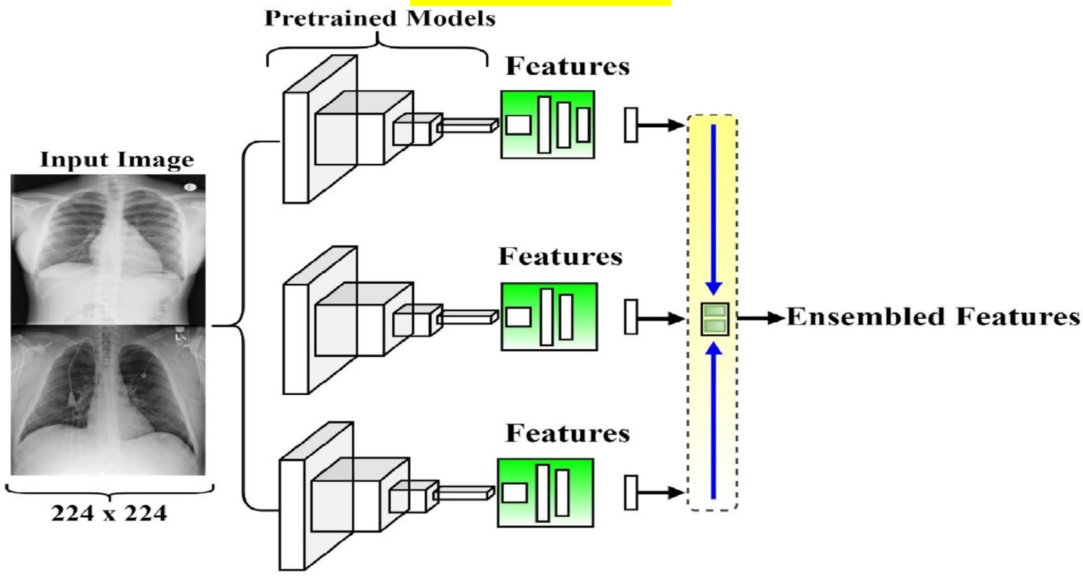


Hybrid deep learning framework

- o Pre-trained ensemble DeepLrn : deep feature extractors
- o Transformer encoder based on the self-attention mechanism + MLP
- \+ Pneumonia accurate identification

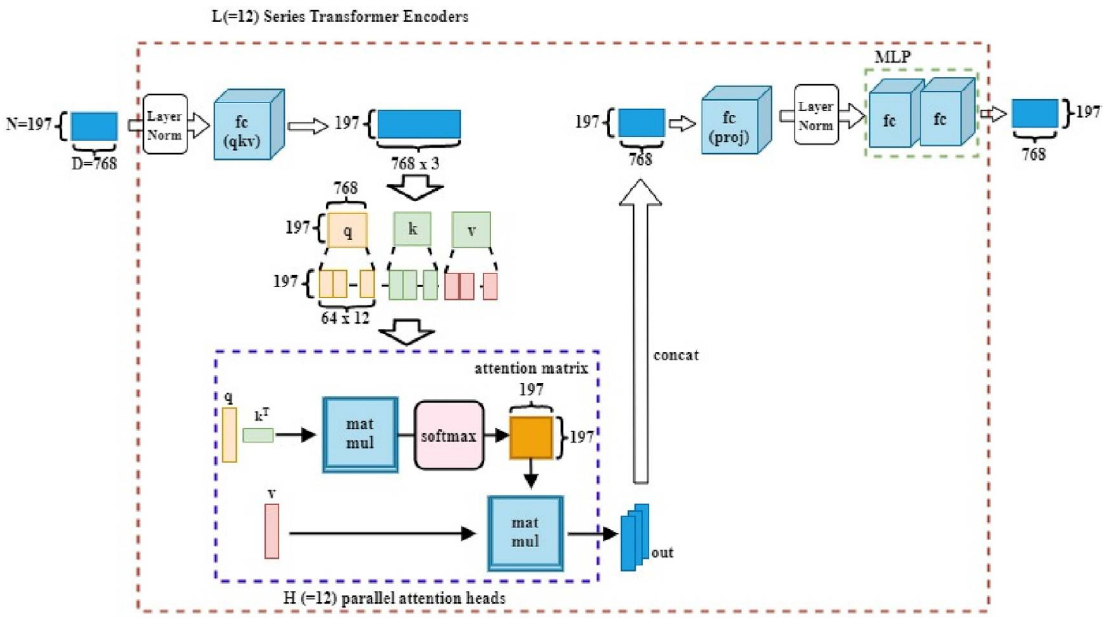
xAI.Med. | 2022-07

Ensemble architecture



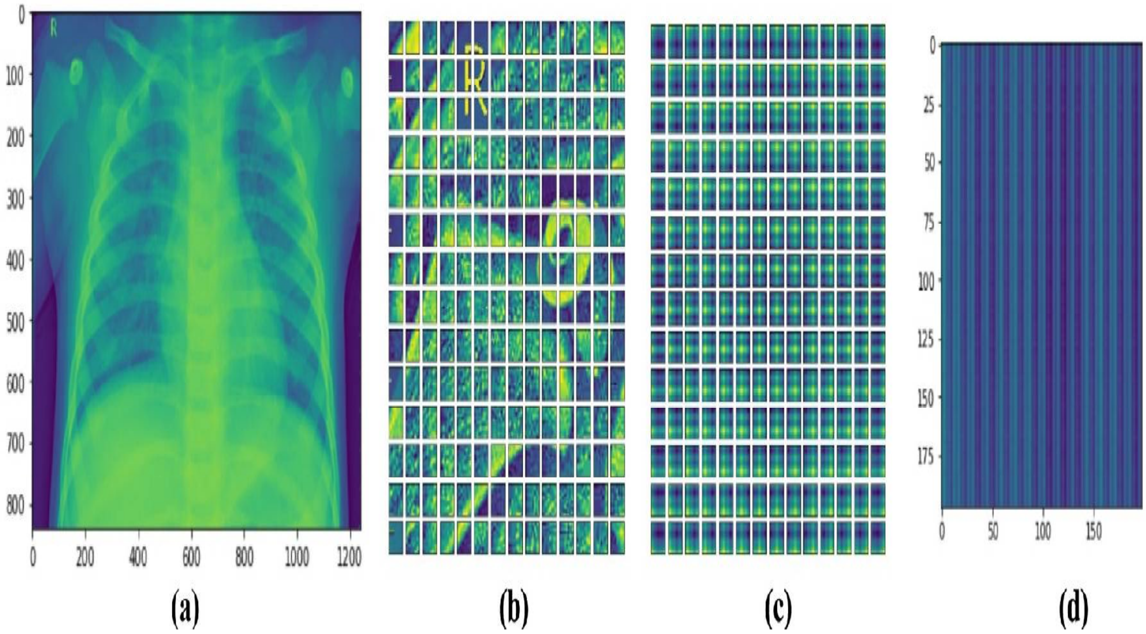
xAI.Med. | 2022-07

Visualization of a Multi-head self-attention Network and MLP blocks



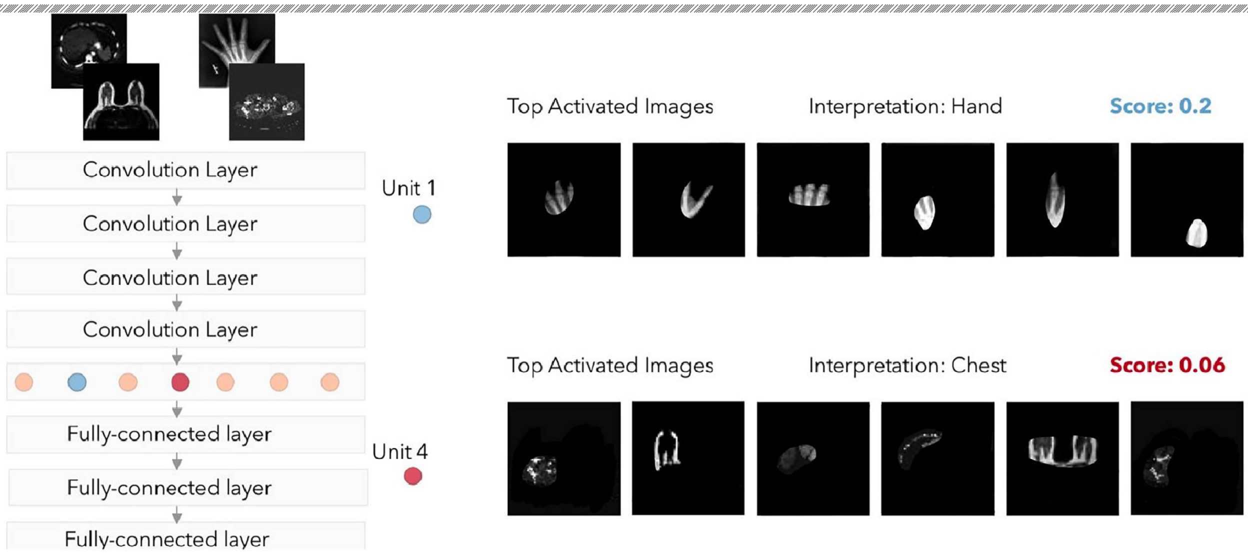
xAI.Med. | 2022-07

Visualization steps of the proposed transformer encoder model



(a) Input :chest X-ray image
(b) Input image divided into equal size nonoverlapping patches
(c) Learnable position embeddings of the input image patches
(d) Demonstrates the corresponding attention matrix

xAI.Med. | **Case Study** | 2023-03

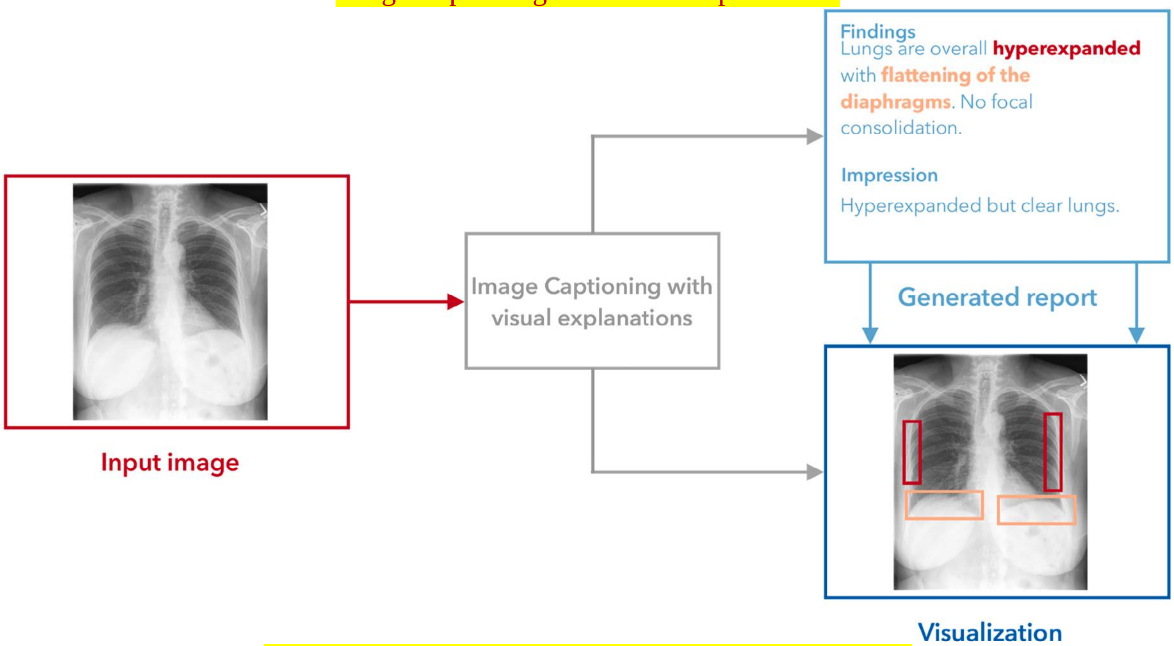


MedNIST dataset

- ✓ Top row represents activations for hands images,
- ✓ Row below contains images of breasts and the abdomen

xAI.Med.		2023-03

Image Captioning with visual explanations



Visualization on the input image using bounding boxes

- ✓ Findings: hyperexpansion and flattened diaphragms

xAI.Med.	t-SNE	2023-03

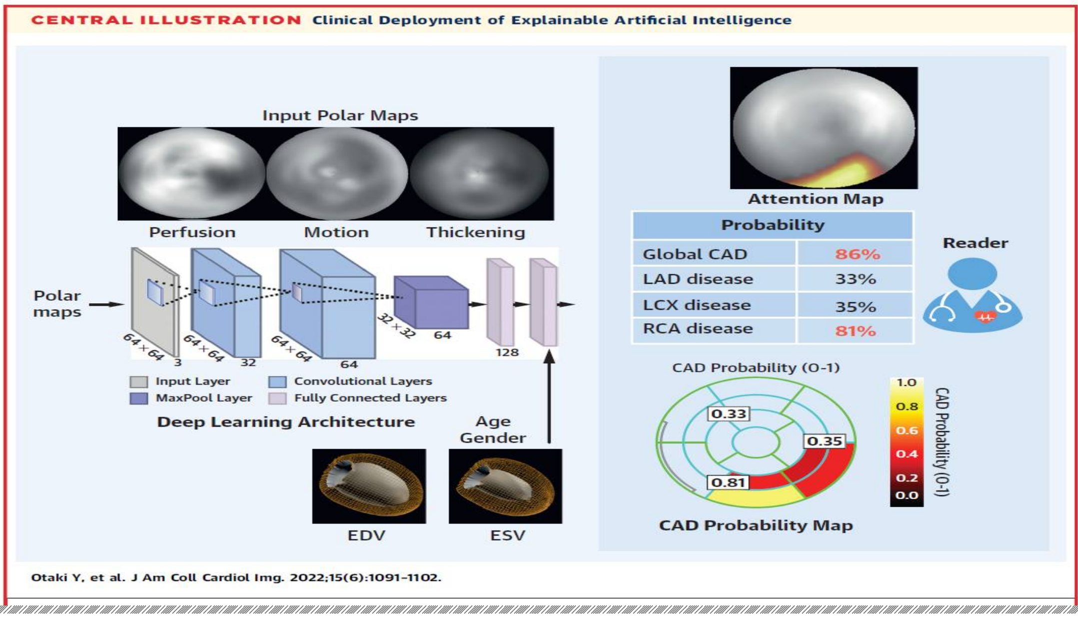
. t-SNE Clustering

- ✓ 'BreastMRI' (0), 'AbdomenCT' (1), 'Hand' (2), and 'CXR' (3).
- ✓ Detected clusters indicate a good separability between the classes

Heart

xAI.Med. 2022-63

Model : CAD-DL



Input:

- ✓ The raw myocardial perfusion, wall motion, thickening images
- ✓ Sex, age, and cardiac volumes

Output

- ✓ Per-vessel probability

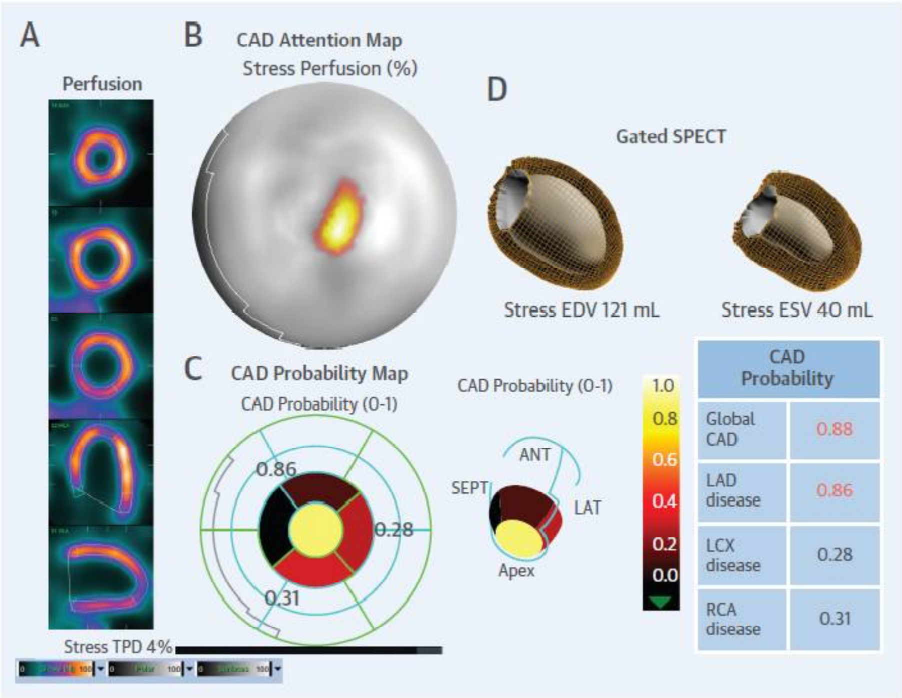
Diagnosis

✓ CAD : coronary artery disease; EDV : end-diastolic volume; ESV : end-systolic volume; LAD : left anterior descending artery; LCX : left circumflex artery; RCA : right coronary artery

xAI.Med. 2022-63

85% stenosis in the proximal left anterior descending (LAD) artery on coronary angiography
72-year-old male

FIGURE 4 Case Example of Solid-State SPECT-MPI



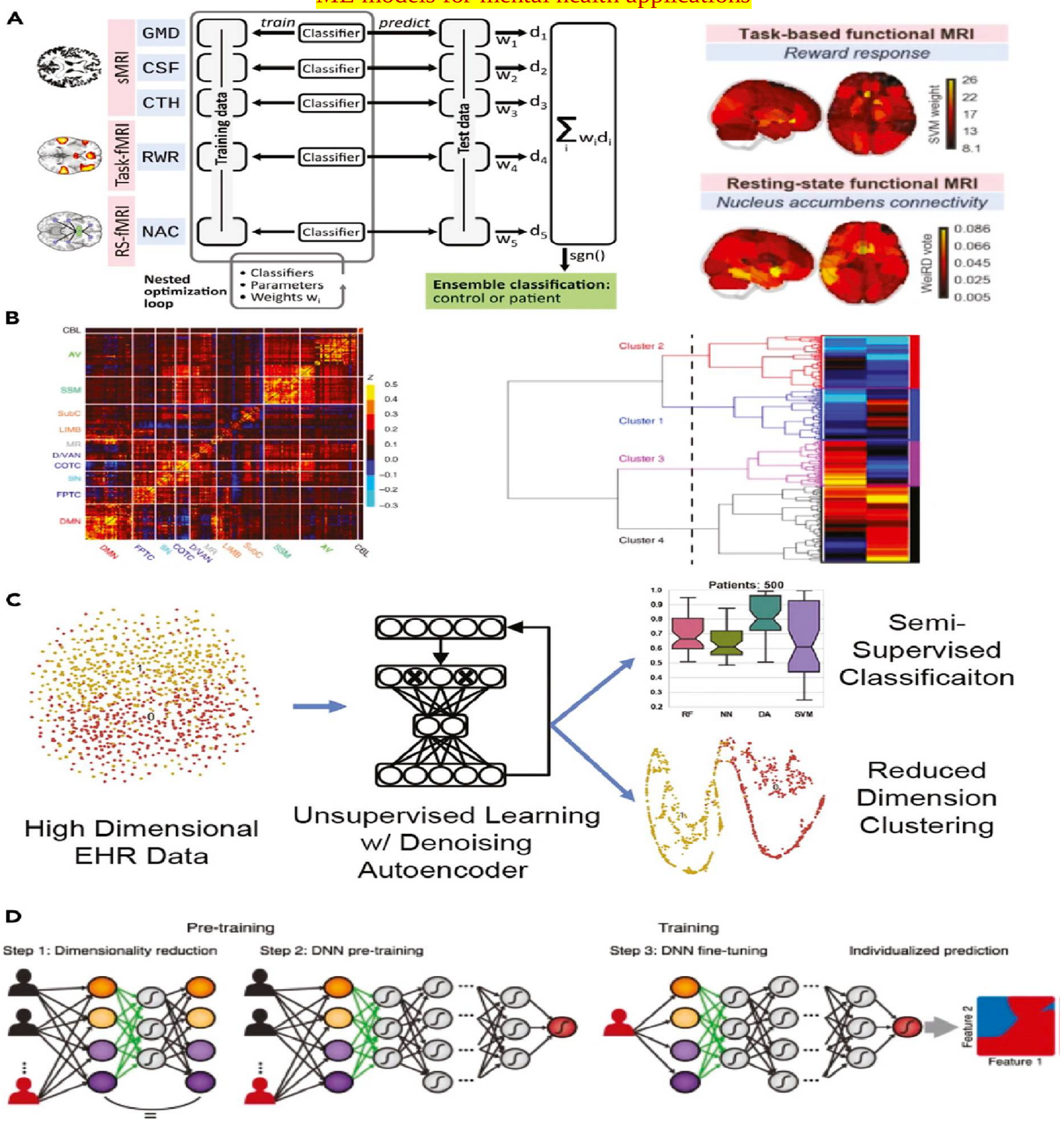
(A) Visual assessment for stress image was interpreted as equivocal (SSS : 2)
(B) CAD attention map highlights the image regions contributing to CAD prediction overlaid onto perfusion polar map
(C) CAD probability map shows a high probability of CAD and specifically LAD disease with the distal anterior and apical segments contributing to the prediction
(D) Left ventricular volumes from gated SPECT

✓ ANT : anterior; EDV : end diastolic volume; ESV : end-systolic volume; LAD : left anterior escending artery; LAT : lateral; LCX : left circumflex artery; MPI : myocardial perfusion imaging; RCA : right coronary artery; SEPT : septal; SPECT : single-photon emission computed tomography;

Brain

xAI.Med. | Case Study | 2022-10

ML models for mental health applications



[$$Scipt

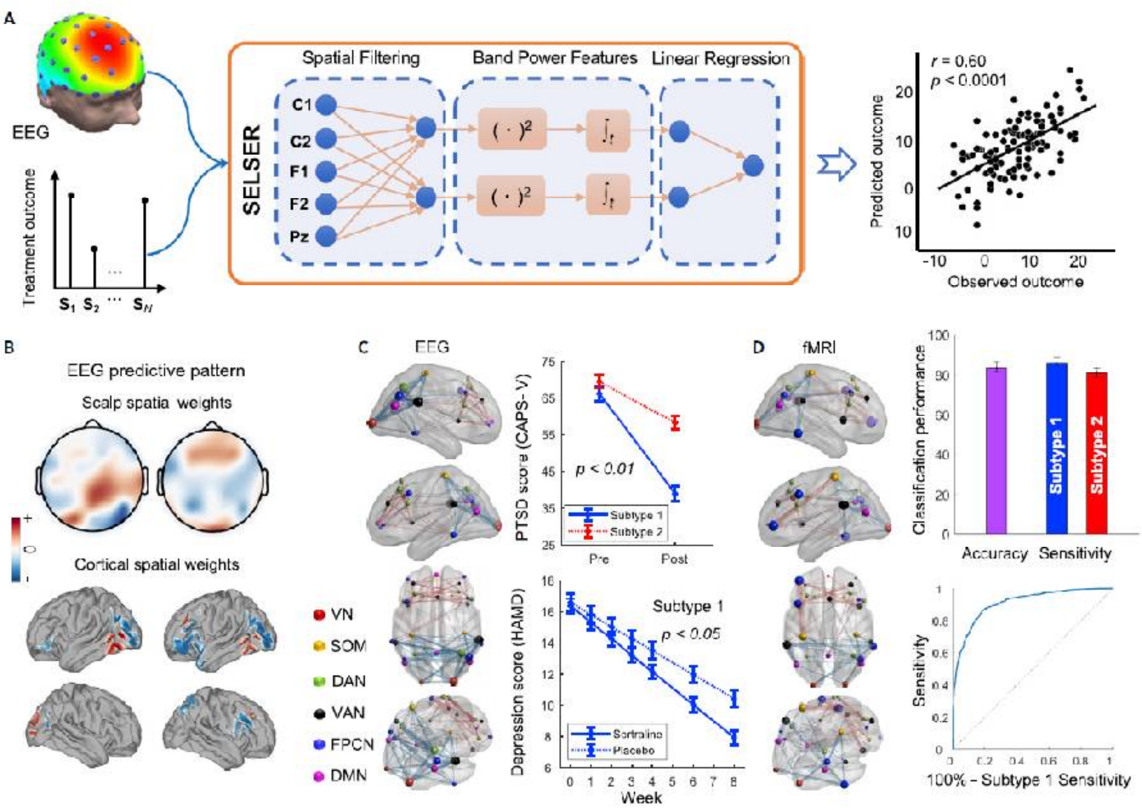
- ✓ (A) "Left: multi-modal supervised classification scheme. Three modality-specific factors are optimized on the training data: classifier types, parameters, and weights. The final diagnostic classificationis based on a weighted sum of decision values, where weights correspond to those estimated during training.
  Right: feature importance maps of functional neuroimaging modalities.

- ✓ (B) Unsupervised learning. Left: whole-brain functional-connectivity matrix averaged across all subjects. z = Fisher-transformed correlation coefficient.
  Right: hierarchical clustering analysis
- ✓ (C) Semi-supervised learning pipeline for phenotype stratification based on EHRs
- ✓ (D) Deep neural networks (DNNs) for group-level and individualized treatment predictions.

Future data points could then be used to forecast symptom onset, treatment response, or other mental health-related variables Scipt$$]

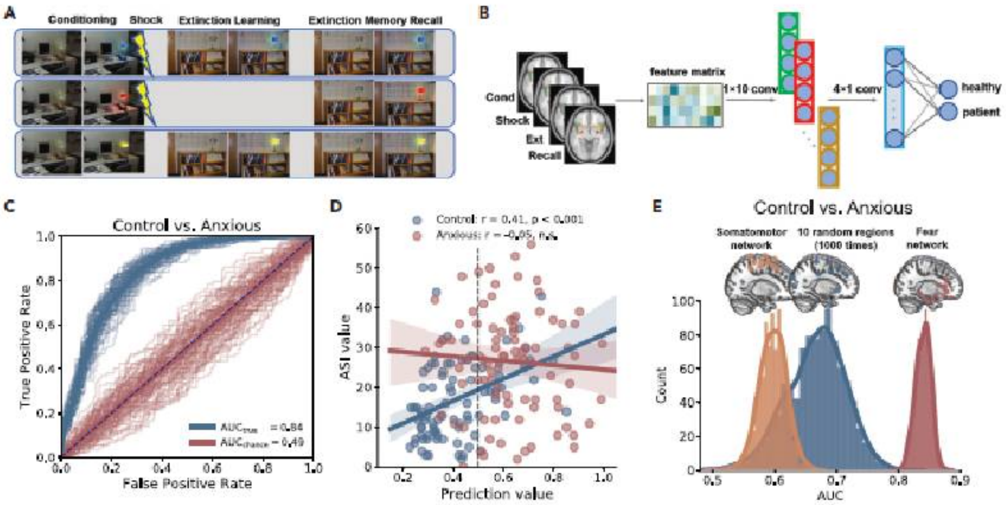
xAI.Med .		2022-10

. Concepts and major findings in case studies 1 and 2



- (A) Illustration of the sparse EEG latent space regression (SELSER) framework in case study 1 for treatment outcome prediction
- (B) Interpretable cortical pattern derived from the scalp pattern
- (C) Distinctive EEG connectivity profiles were identified by sparse K-means for defining psychiatric subtypes in case study 2 on PTSD and MDD. The two
  identified subtypes were further found to predict treatment responsiveness to psychotherapy and antidepressant medication.
- (D) The EEG connectivity-defined subtypes are distinguishable by rs-fMRI connectivity patterns derived from an RVM-based classifier

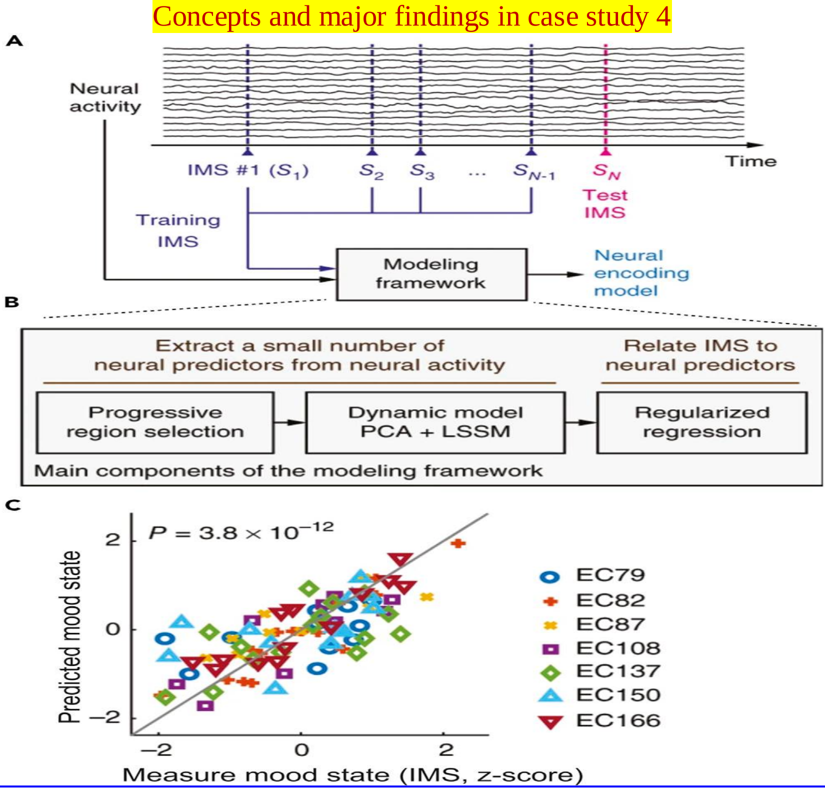
Concepts and major findings in case study



(A) Experimental paradigm.
(B) Schematic of the CNN.
(C) AUC curves produced by CNN versus chance level
(D) The prediction score positively correlated with the anxiety sensitivity index (ASI) for the control group (r = 0.41, p < 0.001) but at the chance level for anxious brains (r = 0.05, p = 0.65)
(E) Distribution of AUCs based on brain activations within 10-node fear randomly selected brain regions

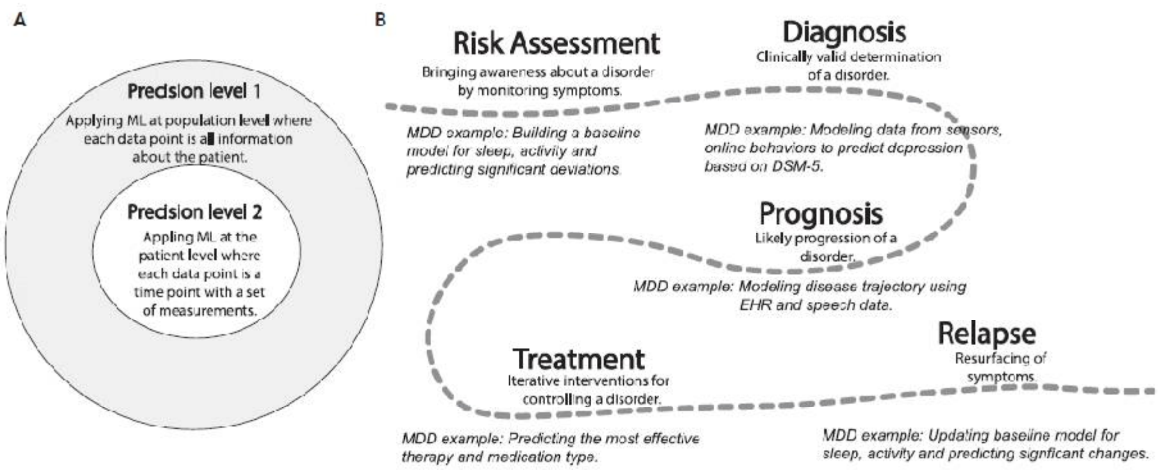
xAI.Med. 2022-10

Concepts and major findings in case study 4



(A) Schematic of cross-validation
(B) Main components of the modeling framework based on both unsupervisedand supervised learning
(C) Cross-validated prediction of the mood state is shown against the truemeasured mood state

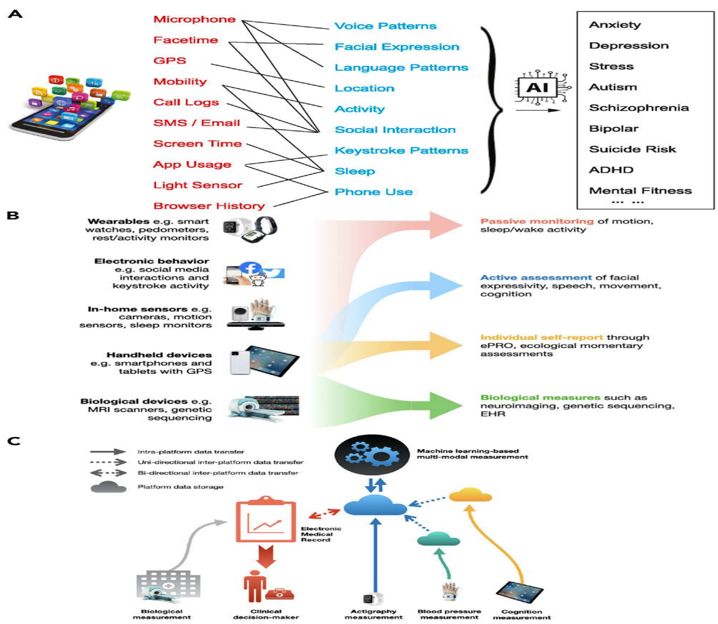
ML for precision psychiatry



(A) Two levels of precision in applying ML for mental health.
(B) Examples of ML applications at various stages of a patient's journey in case of MDD

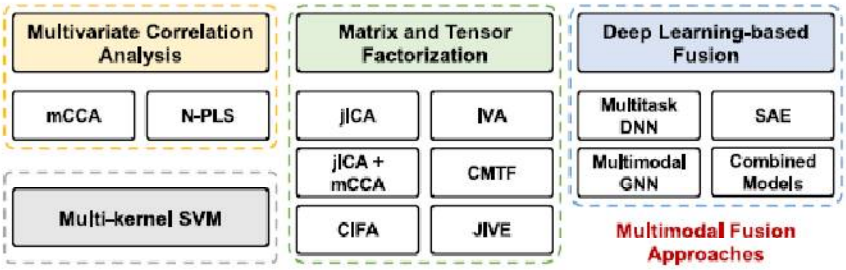
xAI.Med. 2022-10

ML-powered technologies for mental health



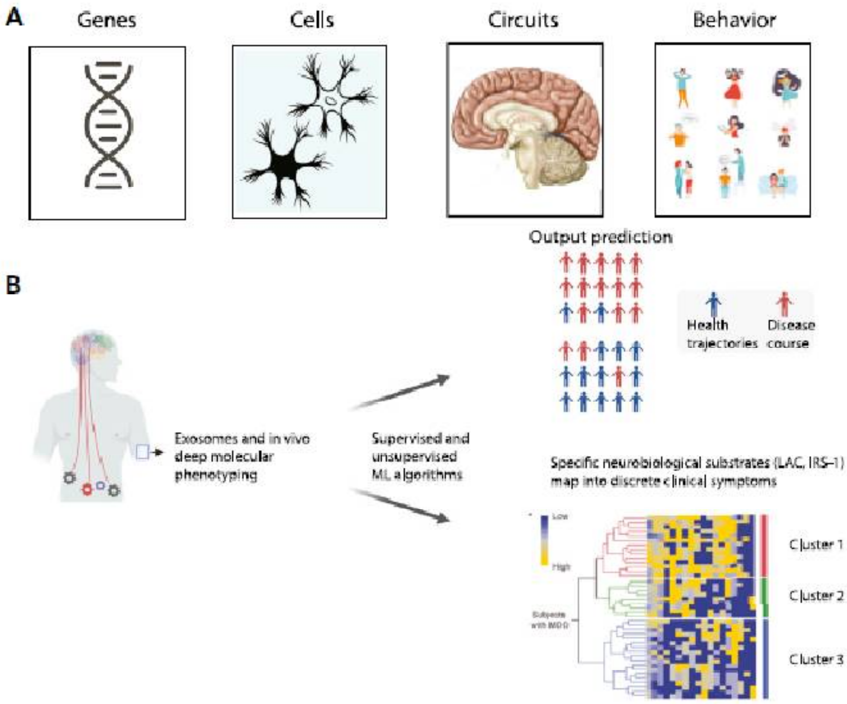
xAI.Med. 2022-10

All compute methods (ACM) for multi-probe data (MPD) fusion fortasks in psychiatry



xAI.Med.	2022-10

ML for molecular phenotyping in psychiatry

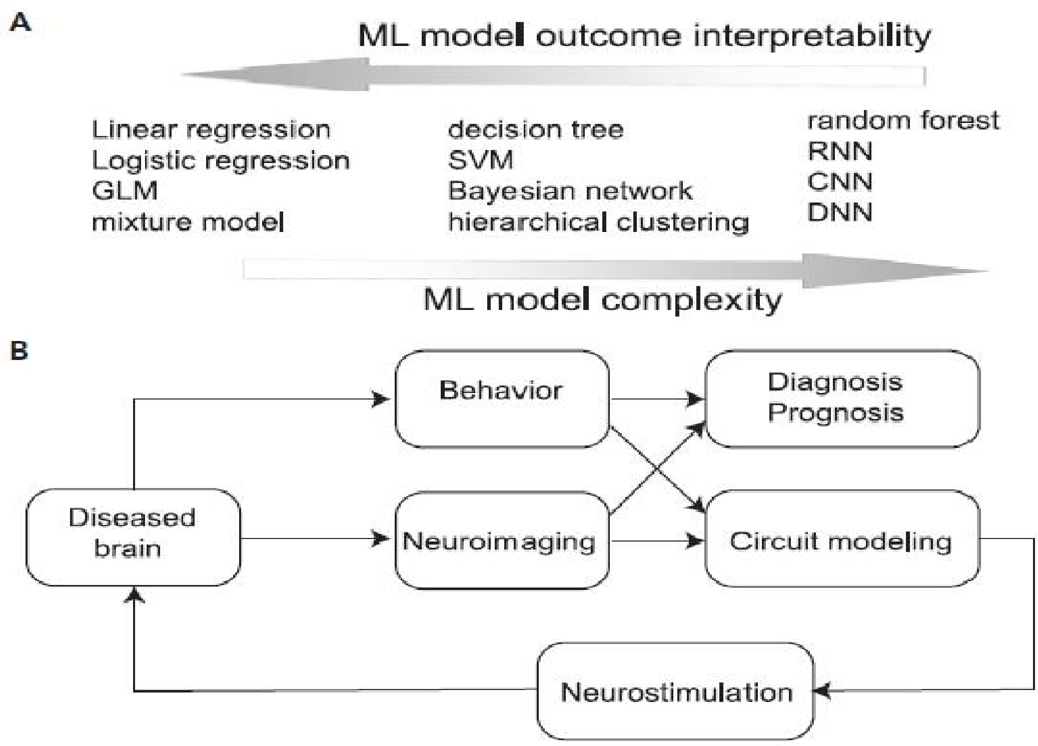


(A) Different levels of interacting variables (genes, cells, circuits) to behaviors in mental illnesses.
(B) Combining ML with novel molecular biology technologies

+ Deep molecular phenotyping of brain plasticity) creates opportunities to develop new mechanistic models for prevention and treatment of clinical endophenotypesof mood and cognitive disorders

xAI.Med.	2022-10

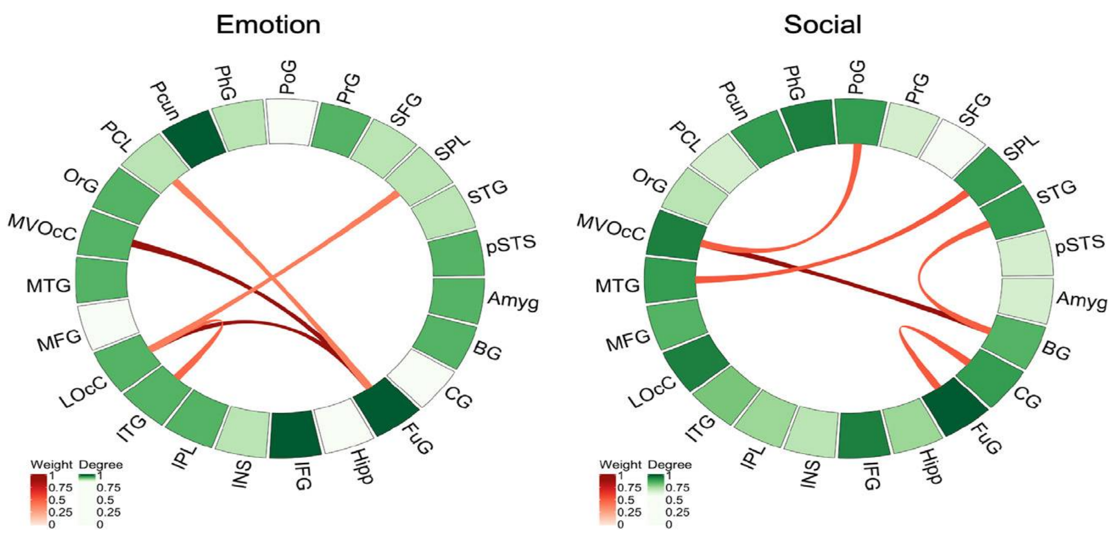
ML model interpretability and closed-loop brain-behavior intervention



(A) A wide spectrum of interpretability in representative ML models.
(B) Schematic of the closed loop of neuroimaging/modeling/neurostimulation

xAI.Med.		2023-12

Circular plots expressing the functional connections between macro-regions in the Emotion (Left) and Social (Right) HCP tasks

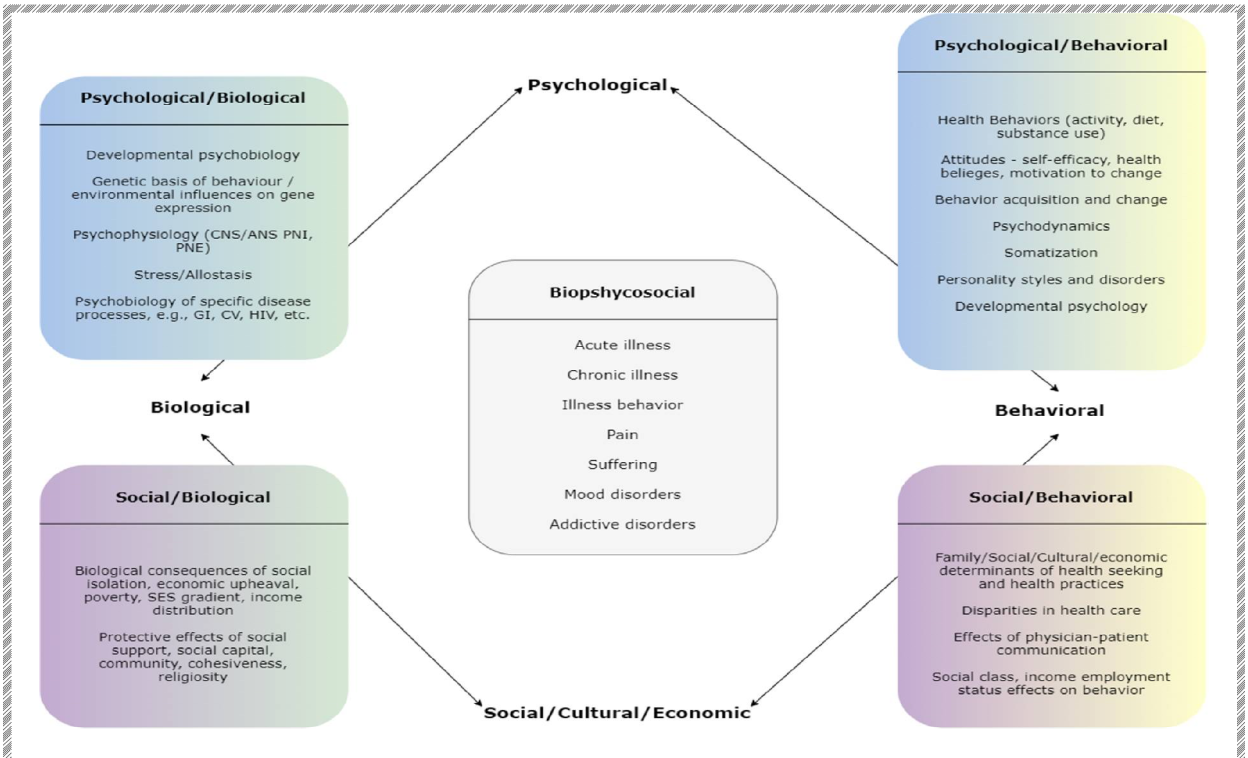


Green shade : node degree
Red shades : connection strength

- ✓ 5 top-most important relevant functional connections displayed

xAI.Med.	**Case Study**	2022-49

. Heterogeneous data in biopsychosocial diagnostic model

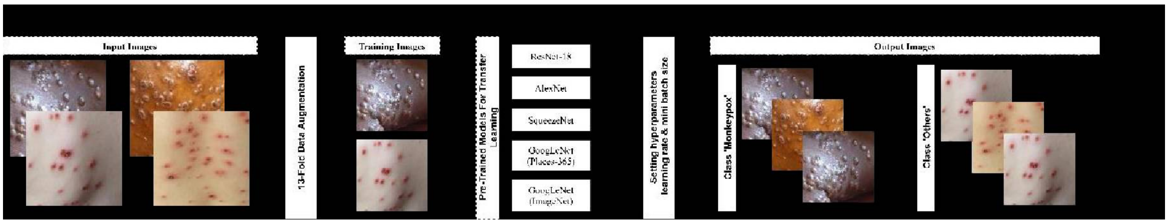


PNI–Psychoneuroimmunology	PNE–Psychoendocrinology
GI–Gastrointestinal	CV–Cardiovascular

Skin

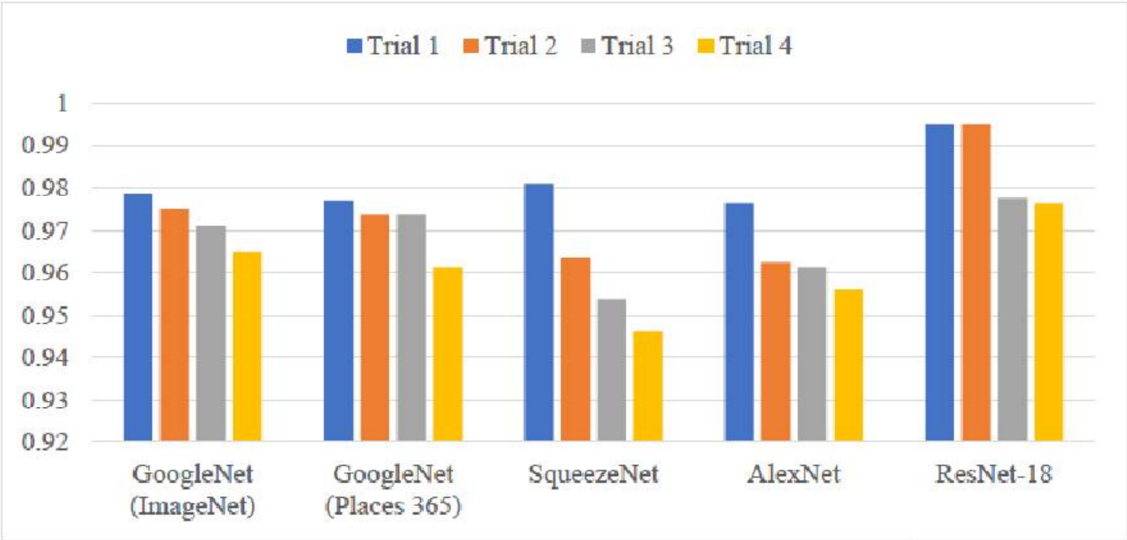
xAI.Med.	Case Study	2023-103

Method Flow diagram to classify monkeypox



xAI.Med.		2023-103

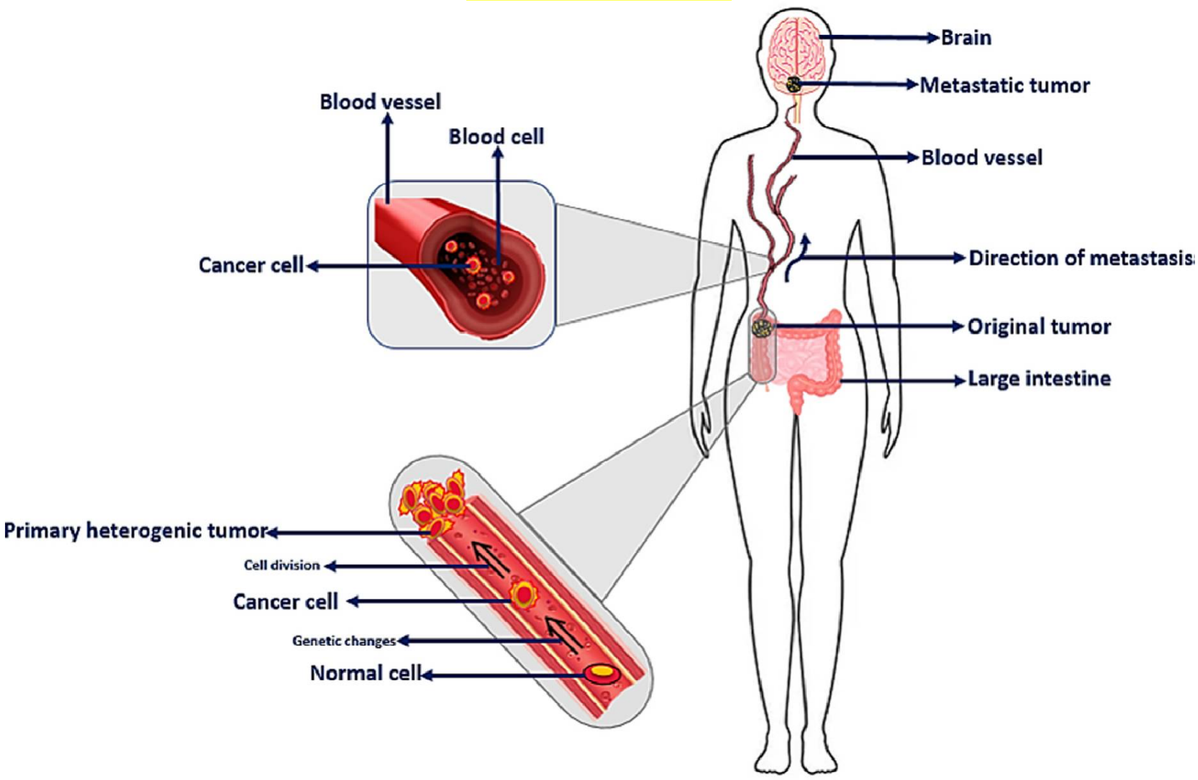
Graphical comparison of the validation accuracies for all four trials



Carcinoma

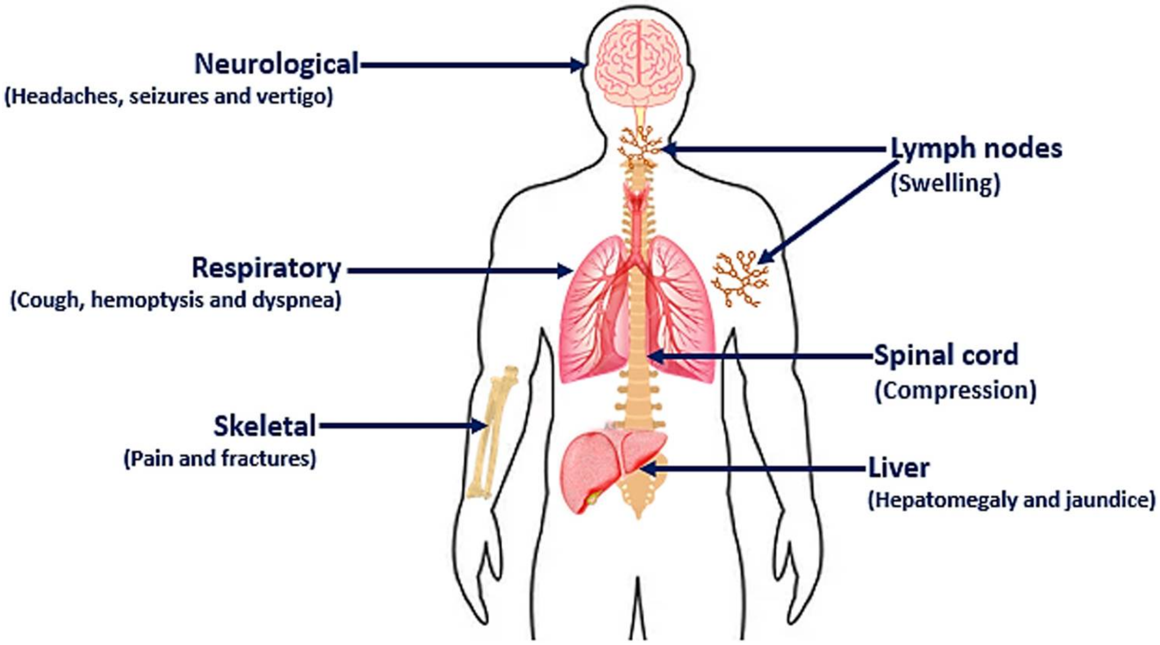
xAI.Med. | 2023-104

Metastasis in the brain



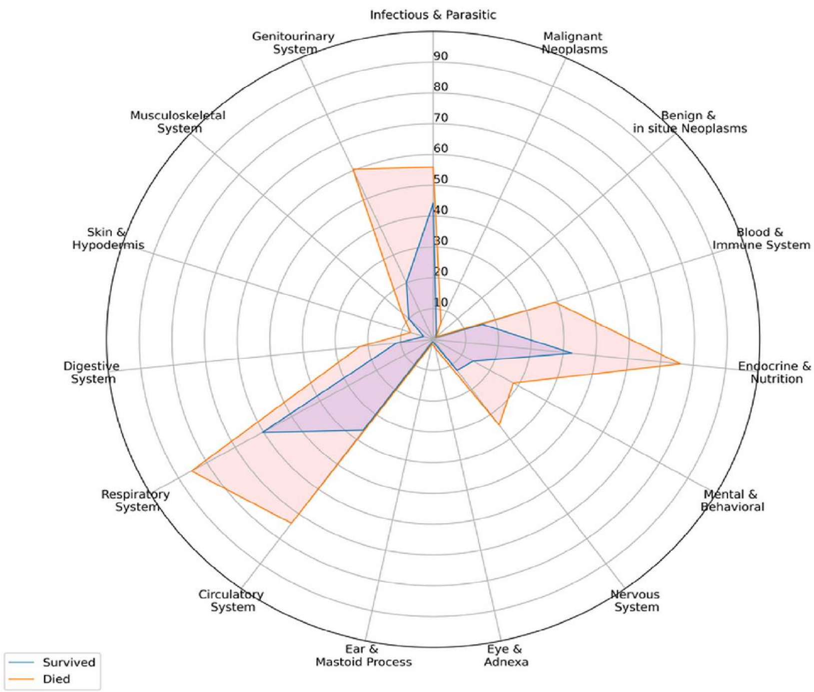
xAI.Med.		2023-104

Common sites and symptoms of metastasis in the body



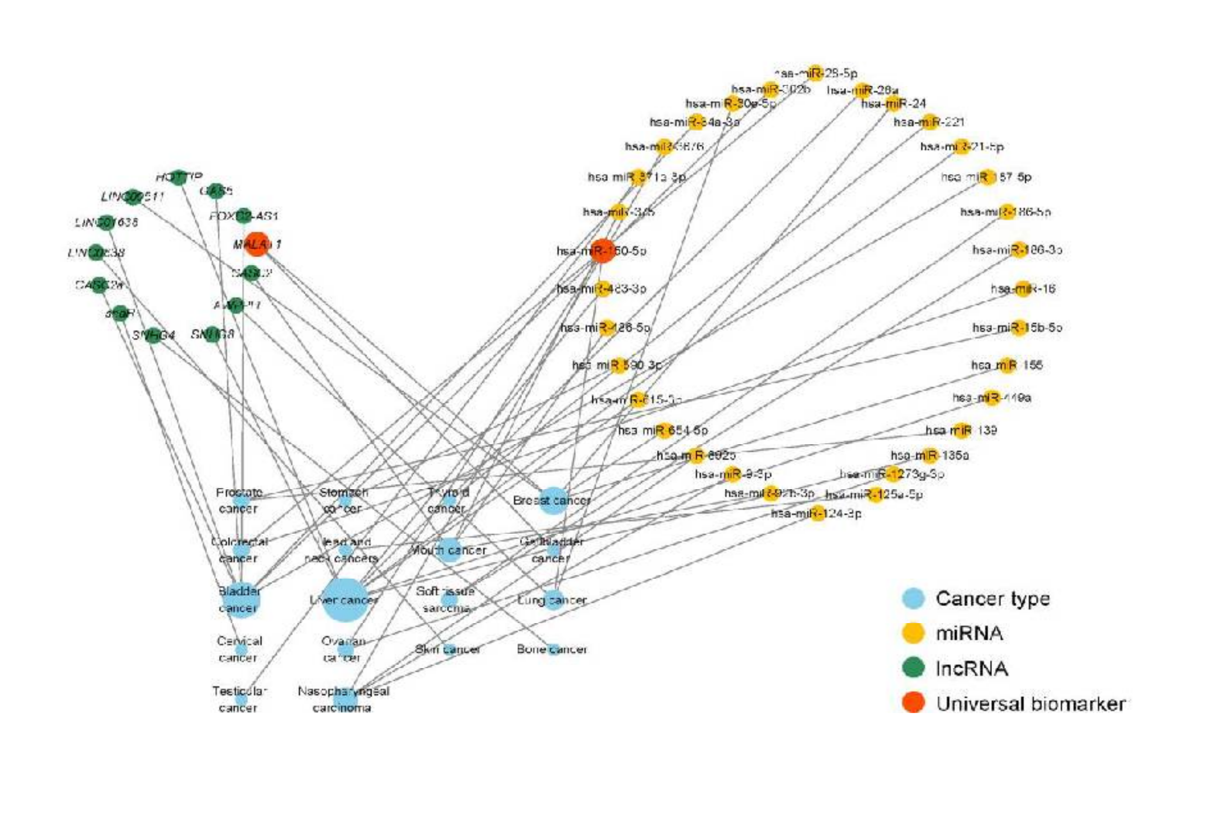
xAI.Med.		2022-27

Proportions of comorbidities by patient outcome

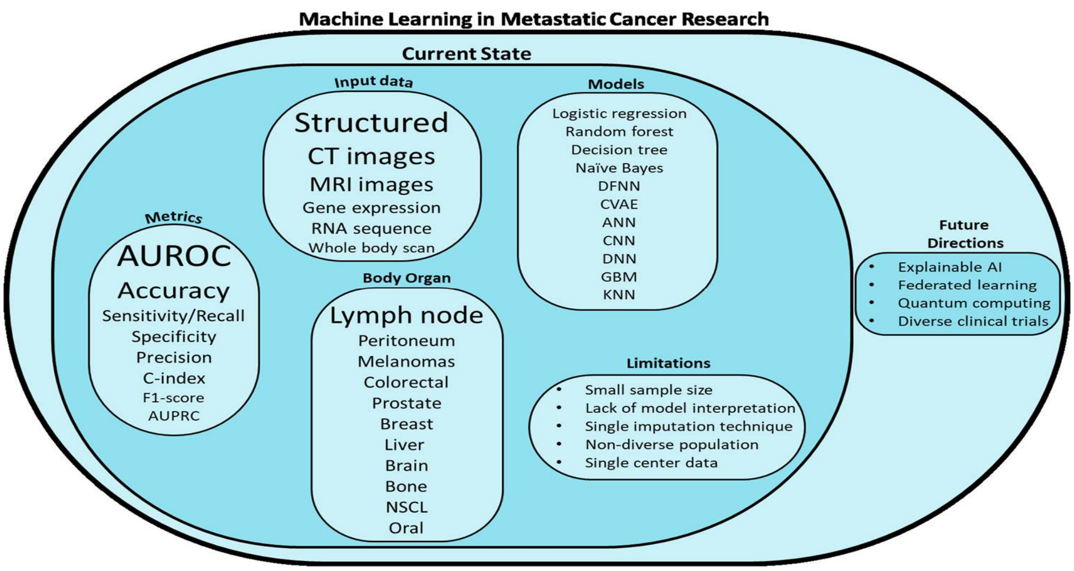


xAI.Med.		2022-11

Cancer

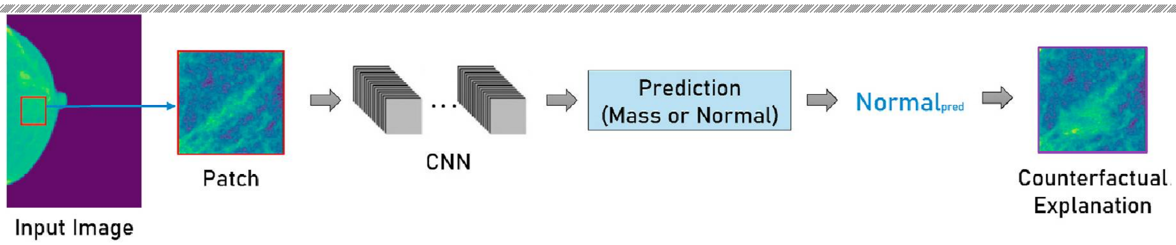


xAI.Med.		2023-104



xAI.Med.		2022-68

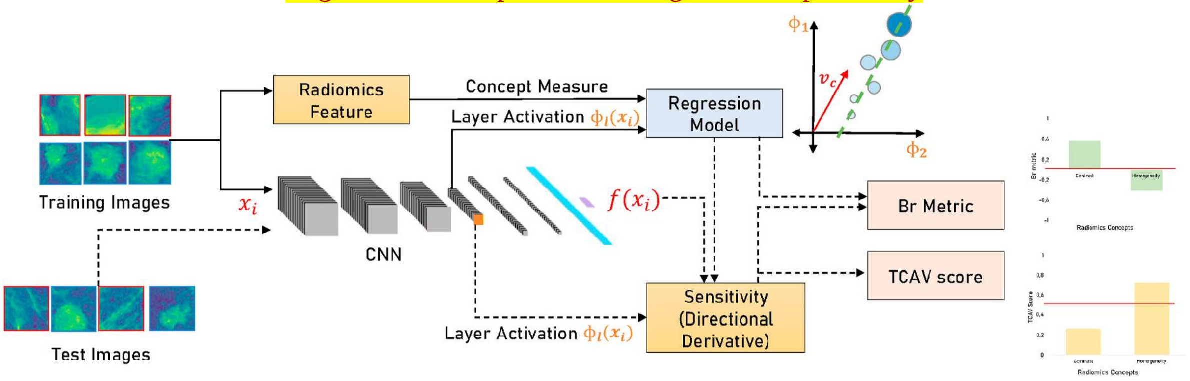
Breast mass prediction in an image patch from a mammogram



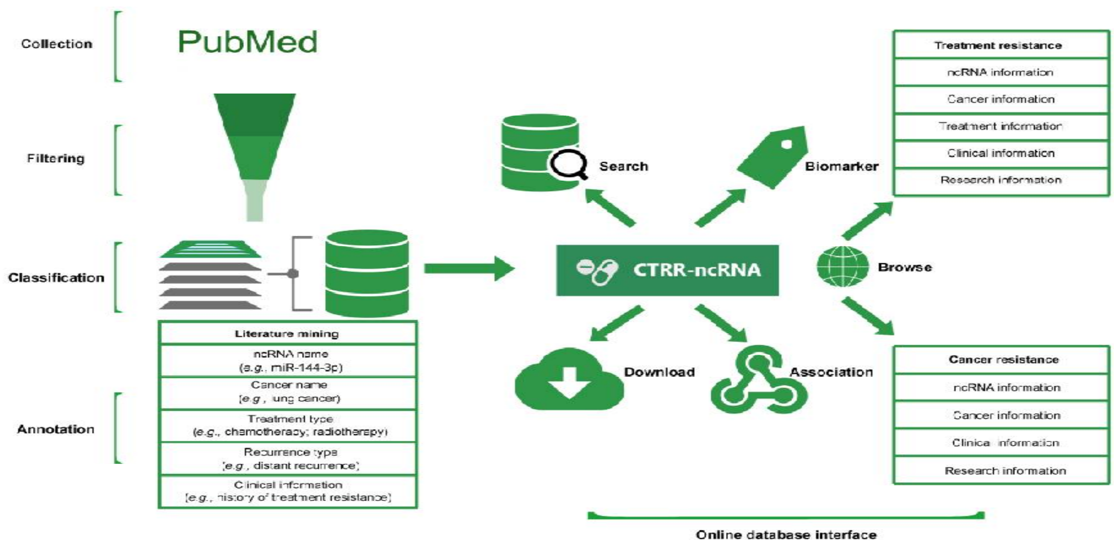
✓ Counterfactual explanation by
- o Applying minimal perturbation to input image to change model's prediction

xAI.Med.	Reg Concept vec	2022-68

Regression Concept Vectors for global interpretability



xAI.Med.	**Case Study**Cancer	2022-11



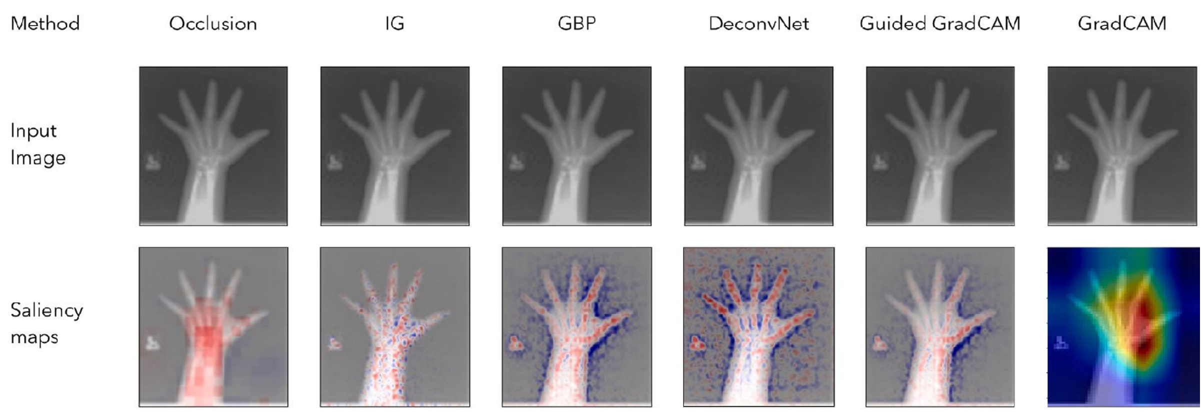
- o Overview of the cancer–ncRNA network in CTR circRNA, circular RNA

(A) ML applications in mobile health
(B) Different types of data collection strategies for digital measurement tools.
(C) A technological infrastructure for the integration of digital measurement tools.

Images

xAI.Med. 2023-22

DenseNet121 trained on MedNIST



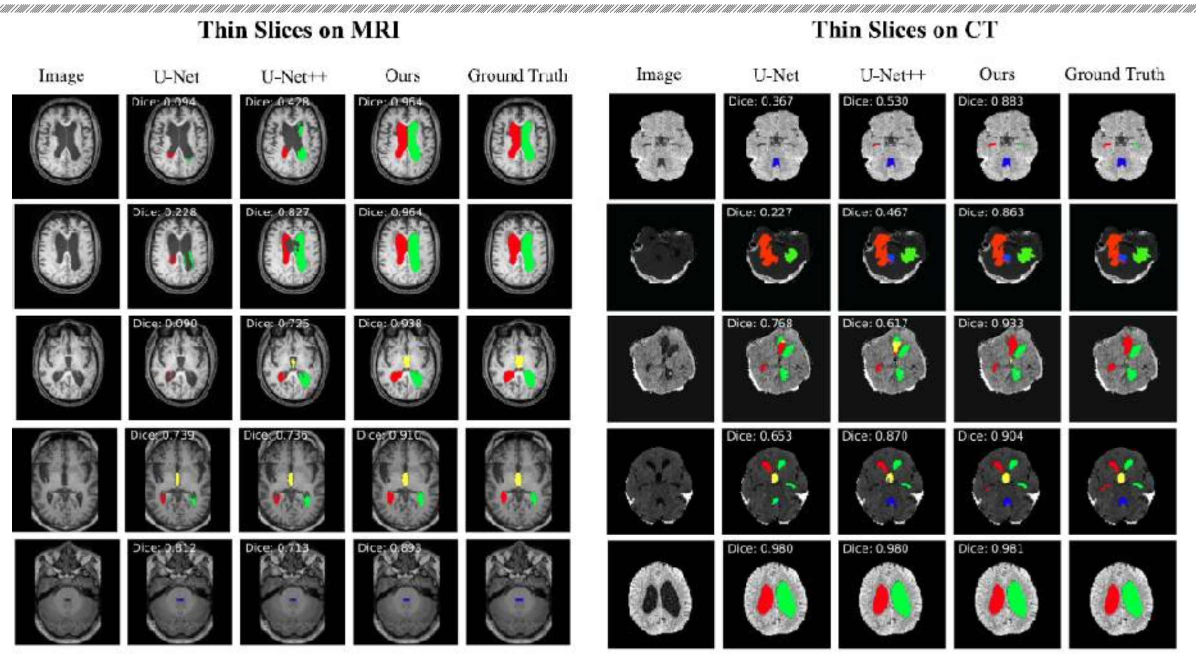
- ✓ Displayed image classified as 'Hand" with a prediction score of 0.98
- ✓ Red areas : positive contribution; blue areas : negative contribution

xAI.Med. 2022-60

Lungs segmentation

xAI.Med. 2022-62

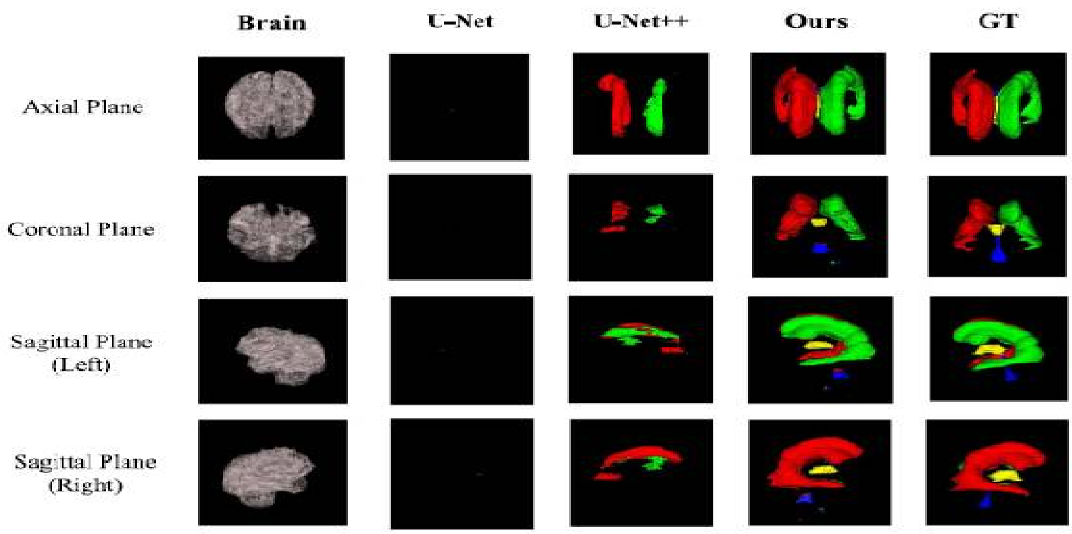
Visualisation of the 3D brain ventricles segmentation



- ✓ Right lateral ventricle is coloured in red;
- ✓ Left lateral ventricle is coloured in green;
- ✓ Yellow coloured region represents the third ventricle;
- ✓ Blue region represents the fourth ventricle

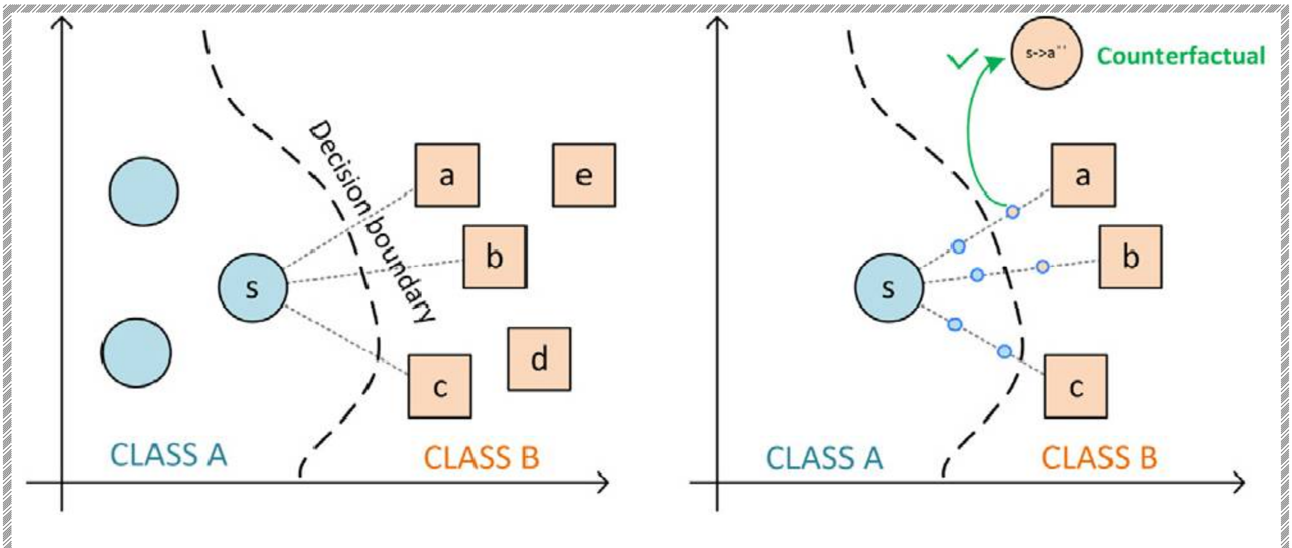
xAI.Med.		2022-

3D-visualisation of predictions on thin-slice MRI images for each ventricle segmented



Case Studies

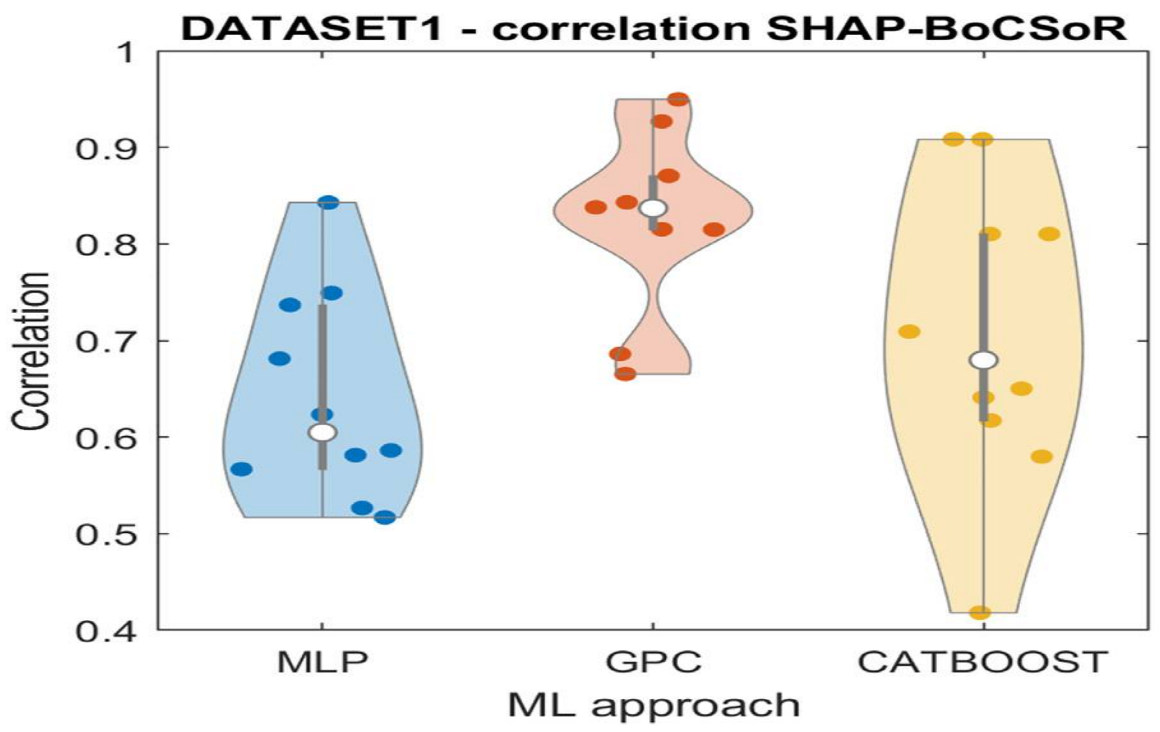
xAI.Med.	Minimally-different counterfactual for instances belonging to class A	2023-12



Alg.

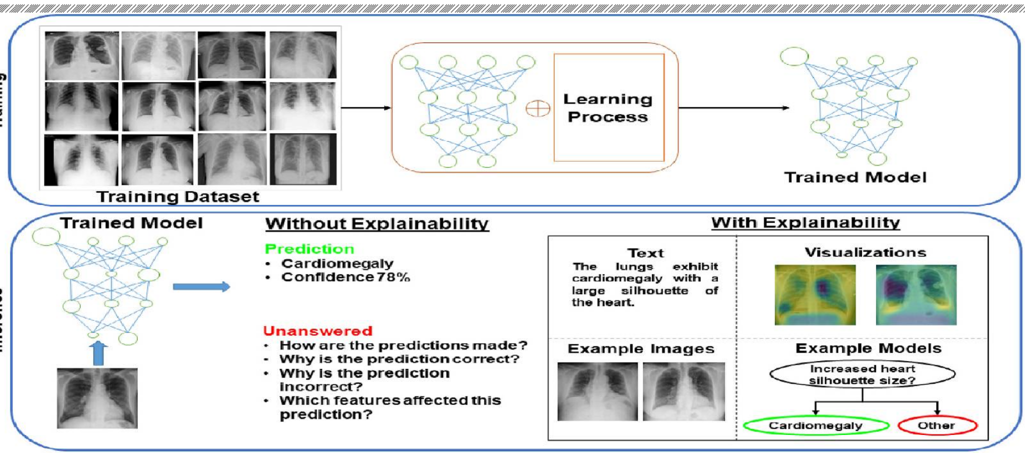
1) The three nearest neighbors of s from the counterfactual class (i.e., class B) considered
2) For each of them, two equally distant midpoints are generated
3) Among all the midpoints corresponding to a classification outcome equal to B, the closest is considered as the minimally-different counterfactual.

xAI.Med. | **Case Study** | 2023-12

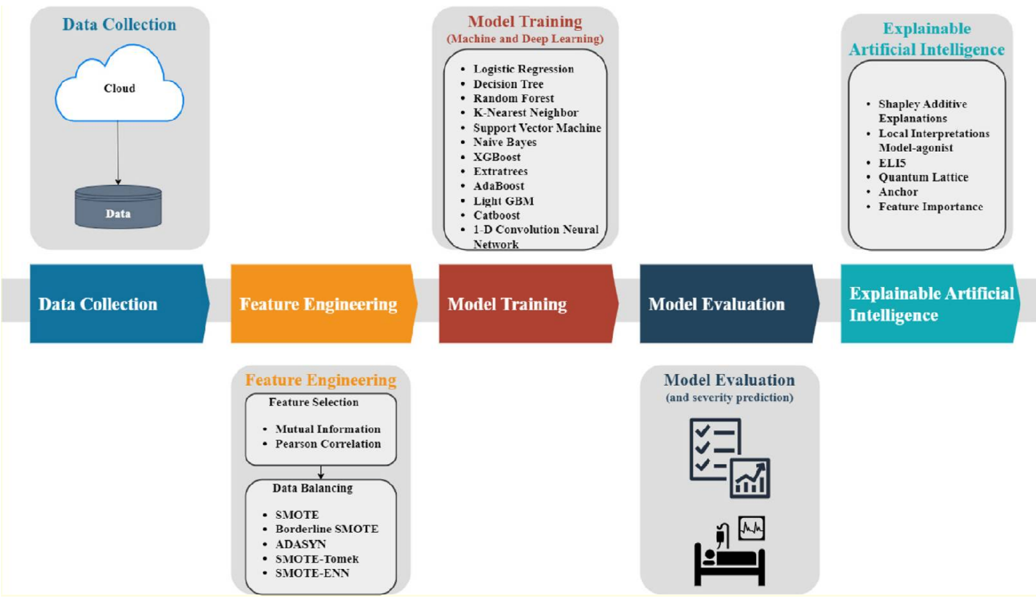


xAI.Med. | Stakeholders**Case Study** | 2023-14

XAI helps stakeholders to understand the model's decision



xAI method pipeline



. ELI5 explanation results

Weight	Feature
0.4228	RESPIRATORY_RATE_MEAN
0.1626	BLOODPRESSURE_DIASTOLIC_MEAN
0.1454	BLOODPRESSURE_SYSTOLIC_MEAN
0.1339	TEMPERATURE_MEAN
0.0478	CALCIUM_MEAN
0.0328	LACTATE_MEAN
0.0229	SAT02_VENOUS_MEAN
0.0149	PH_VENOUS_MEAN
0.0085	BIC_ARTERIAL_MEAN
0.0083	PC02_VENOUS_MEAN

(a)

y=0 (probability **0.913**, score **-2.352**) top features

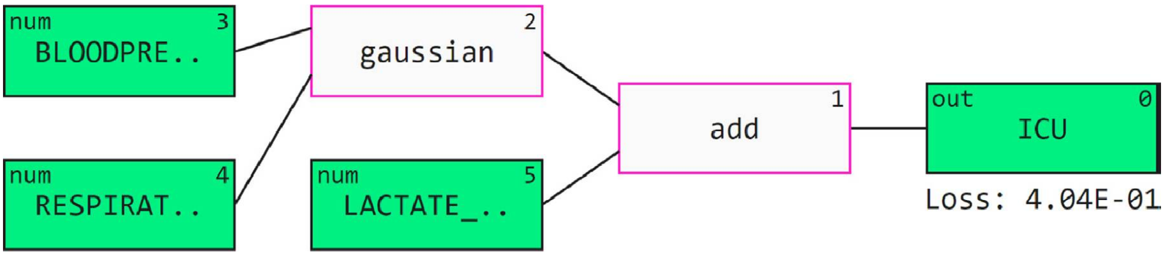
Contribution?	Feature	Value
+1.496	TEMPERATURE_MEAN	2.282
+0.793	CALCIUM_MEAN	0.137
+0.443	BLOODPRESSURE_DIASTOLIC_MEAN	0.780
+0.230	PC02_VENOUS_MEAN	0.008
+0.195	SAT02_VENOUS_MEAN	0.029
+0.136	RESPIRATORY_RATE_MEAN	0.035
+0.095	BIC_ARTERIAL_MEAN	-0.040
+0.059	PH_VENOUS_MEAN	-0.027
-0.159	BLOODPRESSURE_SYSTOLIC_MEAN	0.529
-0.340	<BIAS>	1.000
-0.598	LACTATE_MEAN	-2.303

(b)

✓ Global explanation	✓ Local explanation

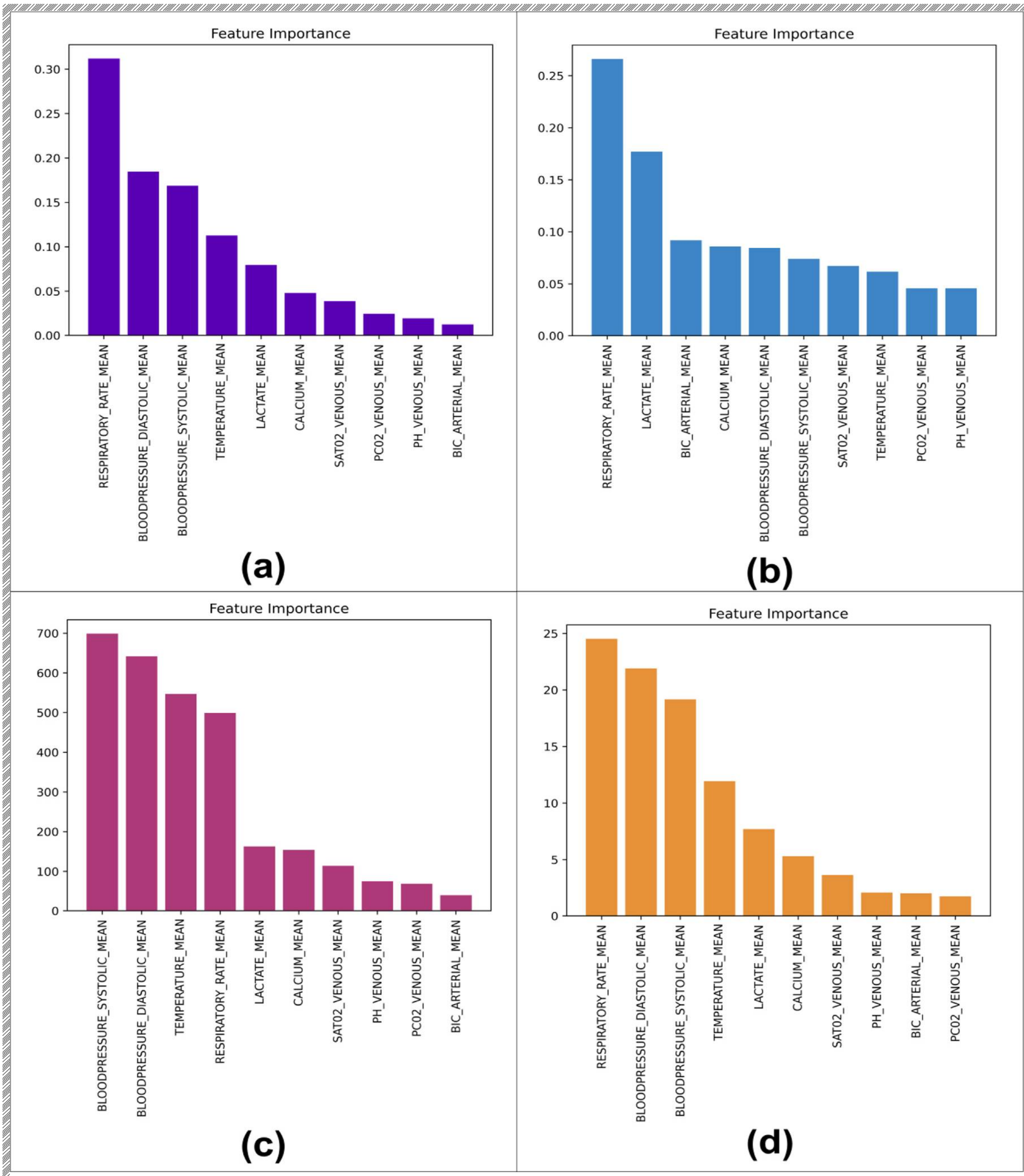
xAI.Med.	2023-17

QGraph plot.



xAI.Med.	2023-17

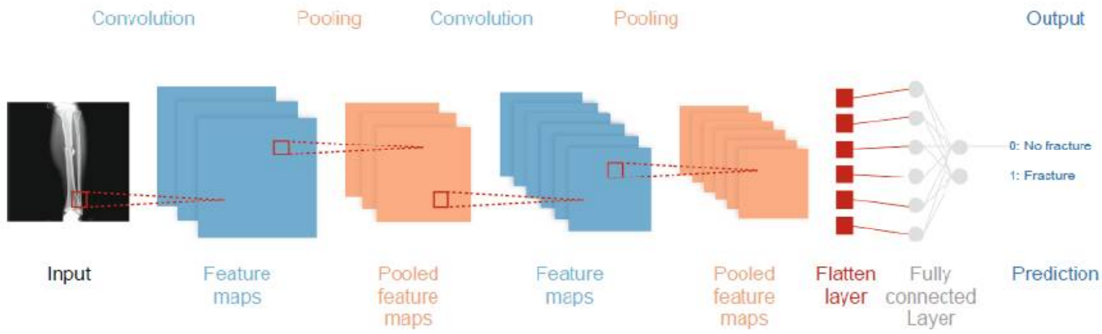
Feature importance plots -- Tree-based Models



(a) Random forest, (b) Xgboost, (c) Light GBM and (d) Catboost

xAI.Med. | 2023-22

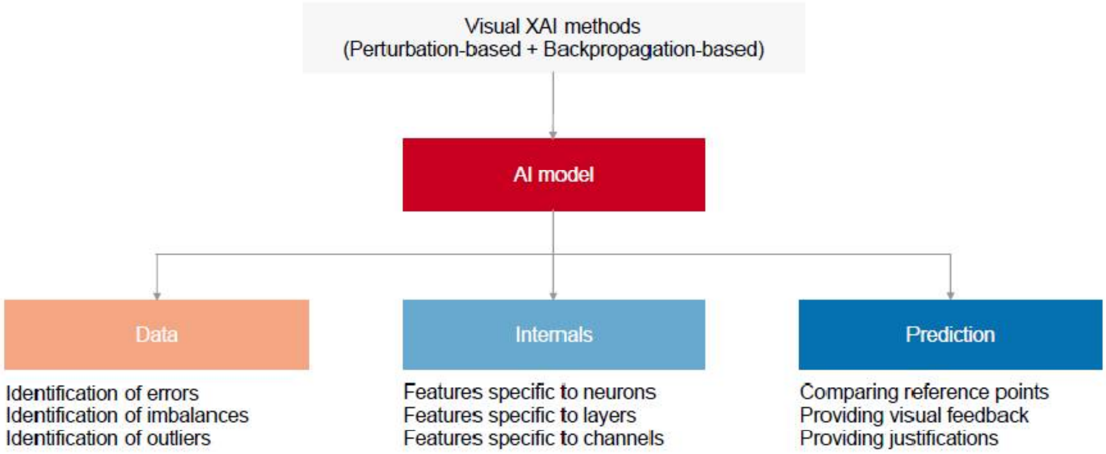
CNN



+ Filters (red rectangles with dashed lines)
+ Pooling layers : to reduce the feature maps' dimensionality
+ Specific filter values are learned during model training
  often called the weights of the CNN

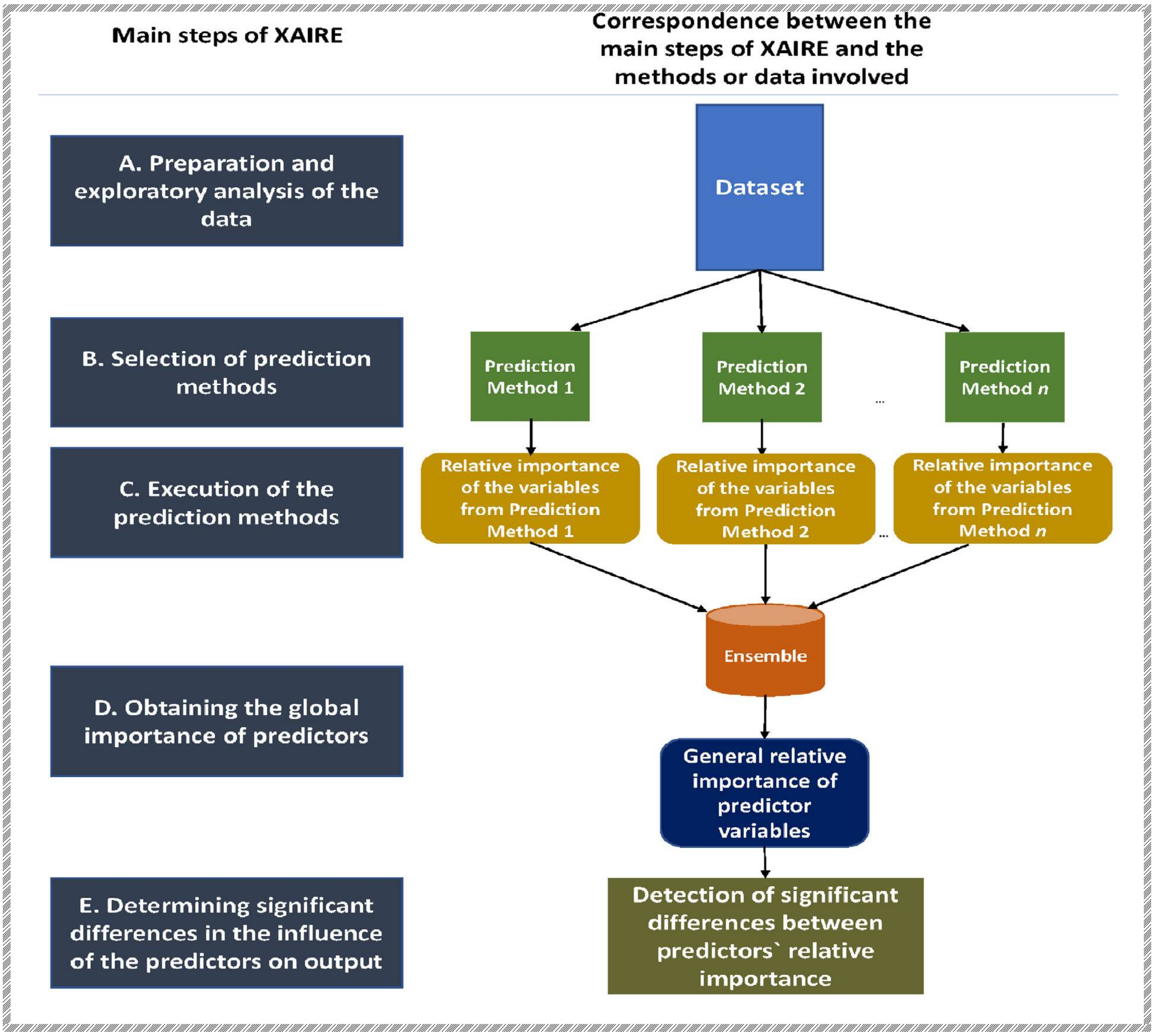
xAI.Med. | 2023-22

Applying visual XAI methods to an AI model
enables the explanation of data, model internals, or model prediction



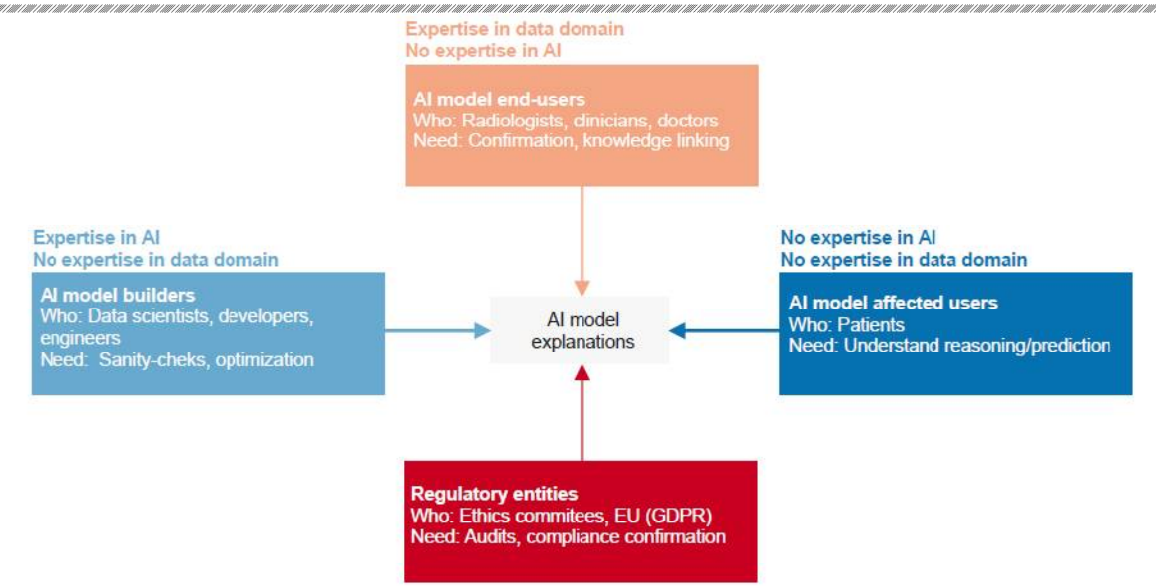
xAI.Med. | **Case Study XAIRE** | 2023-45

XAIRE methodology
for extracting relative importance of features



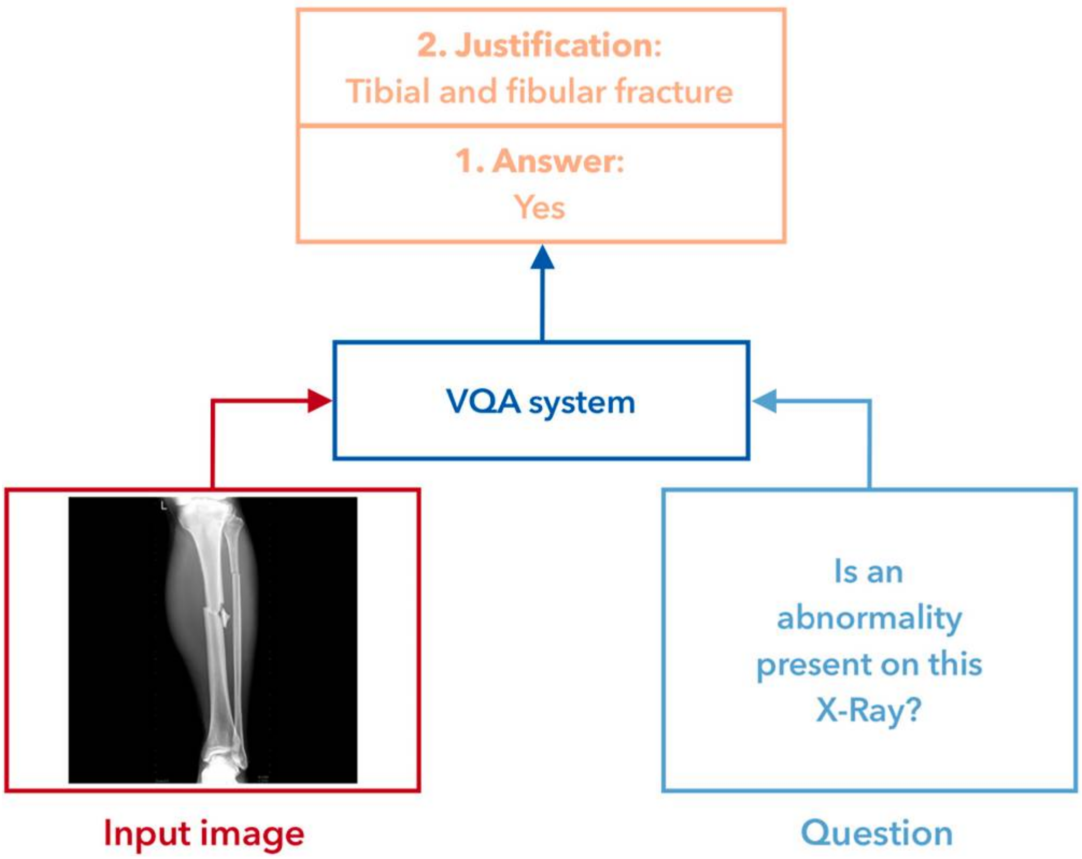
xAI.Med.		2023-22

Aspirations of stakeholder groups
for XAI tools in
AI models
[medical domain]



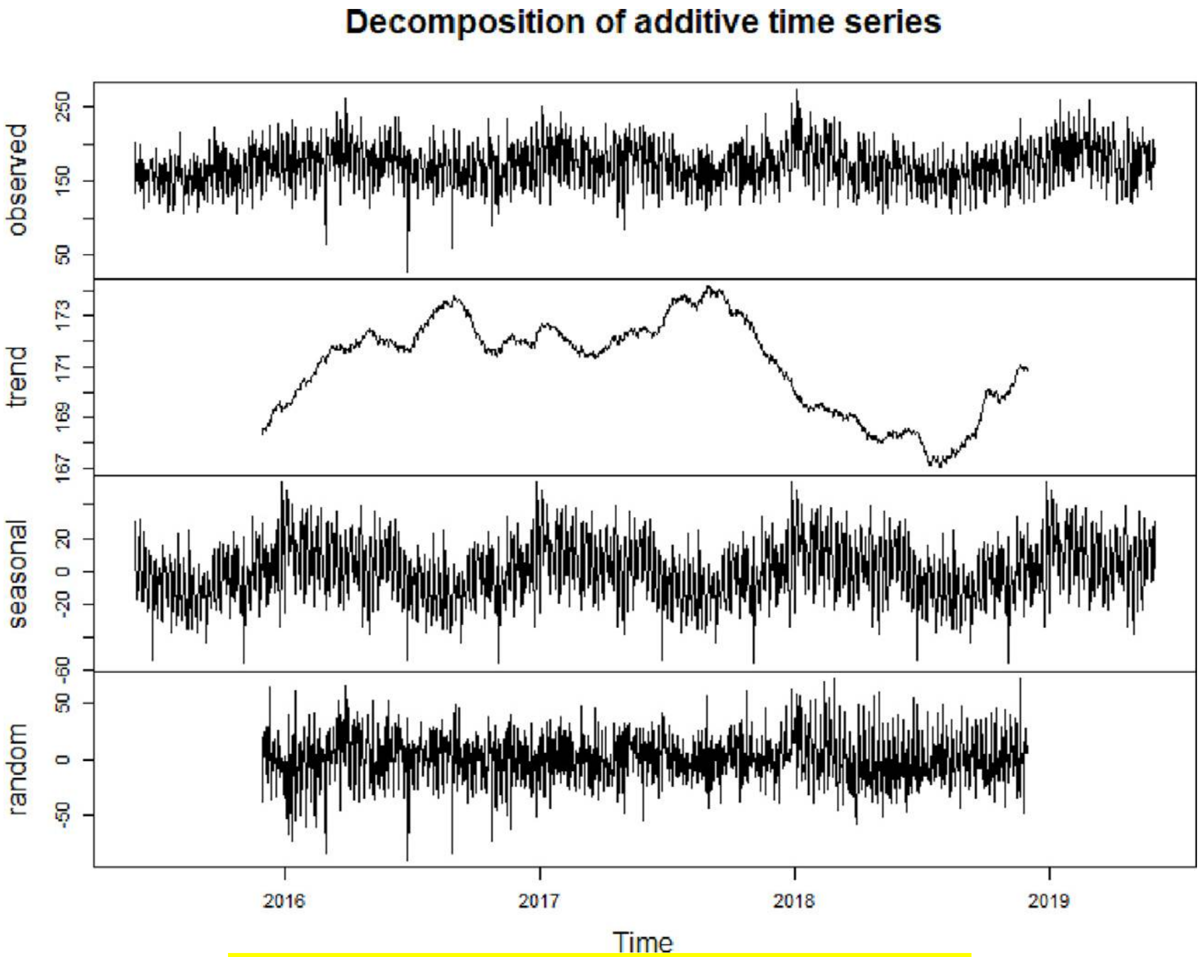
Different explanation needs based on domain knowledge and expertise

Medical imaging domain
VQA system



Miscellaneous

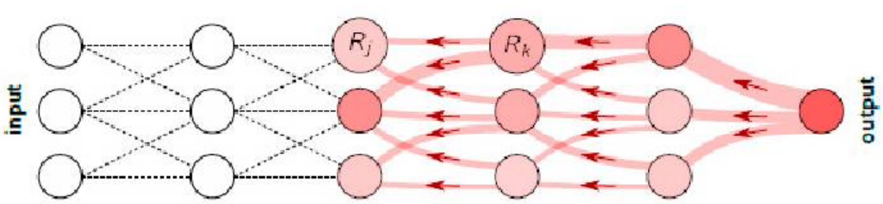
xAI.Med. 2023-45



Additive decomposition of sequential (time) series components

xAI.Med. **Case Study** 2022-24

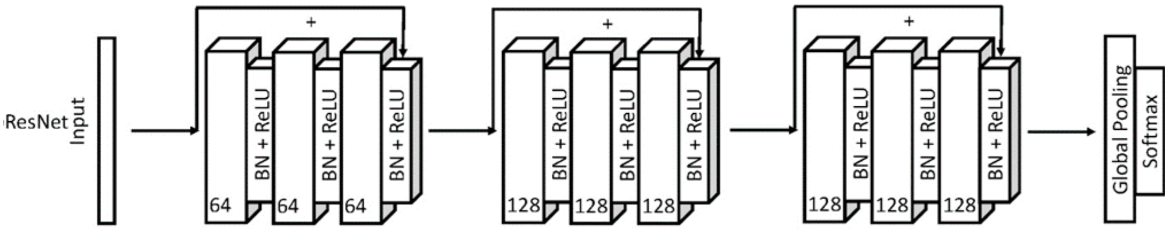
LRP procedure



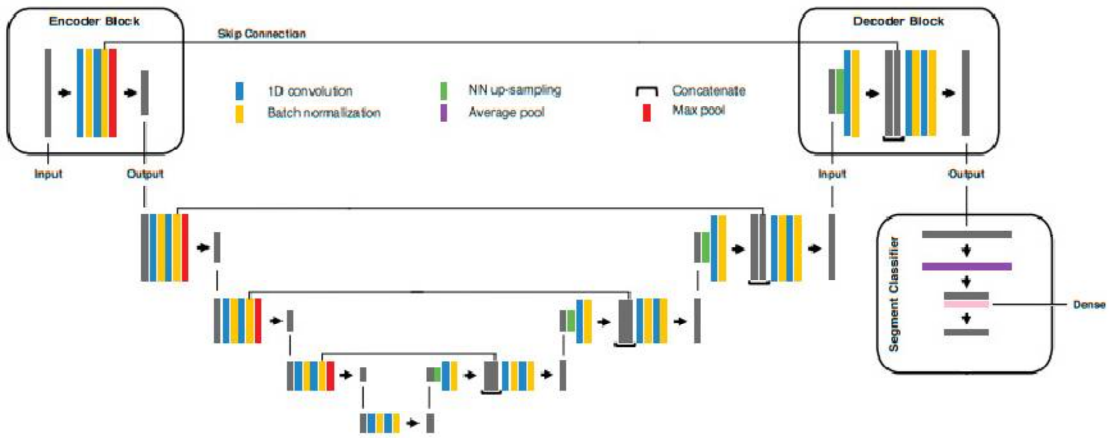
! Each neuron redistributes tolower level as much as it received from higher level

xAI.Med. 2022-24

Modified U-Time architecture

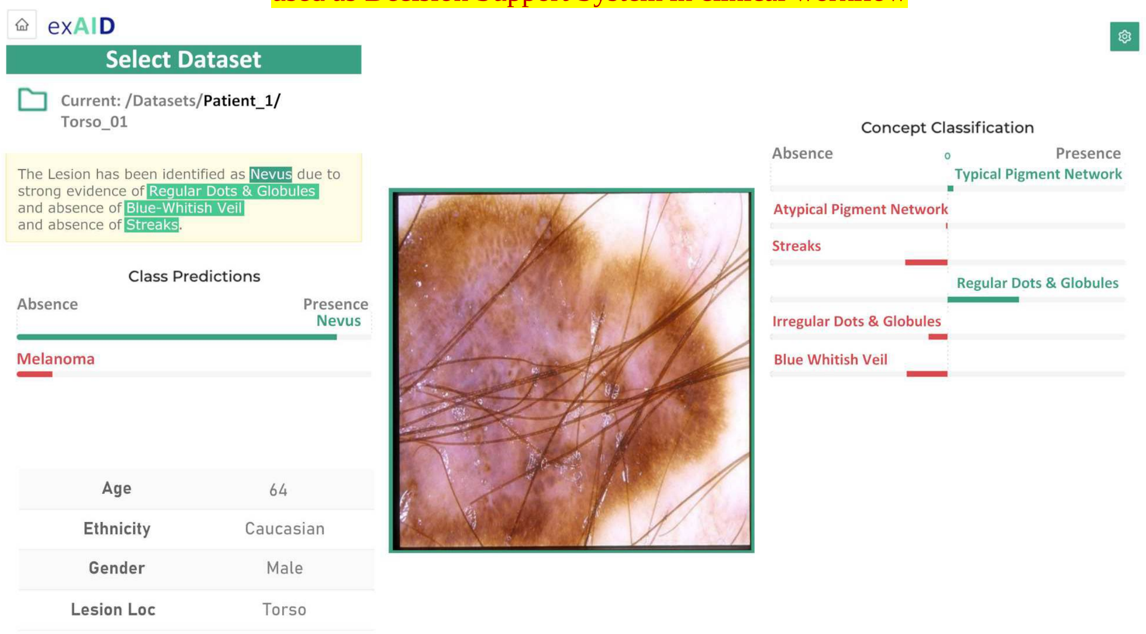


ResNet architecture

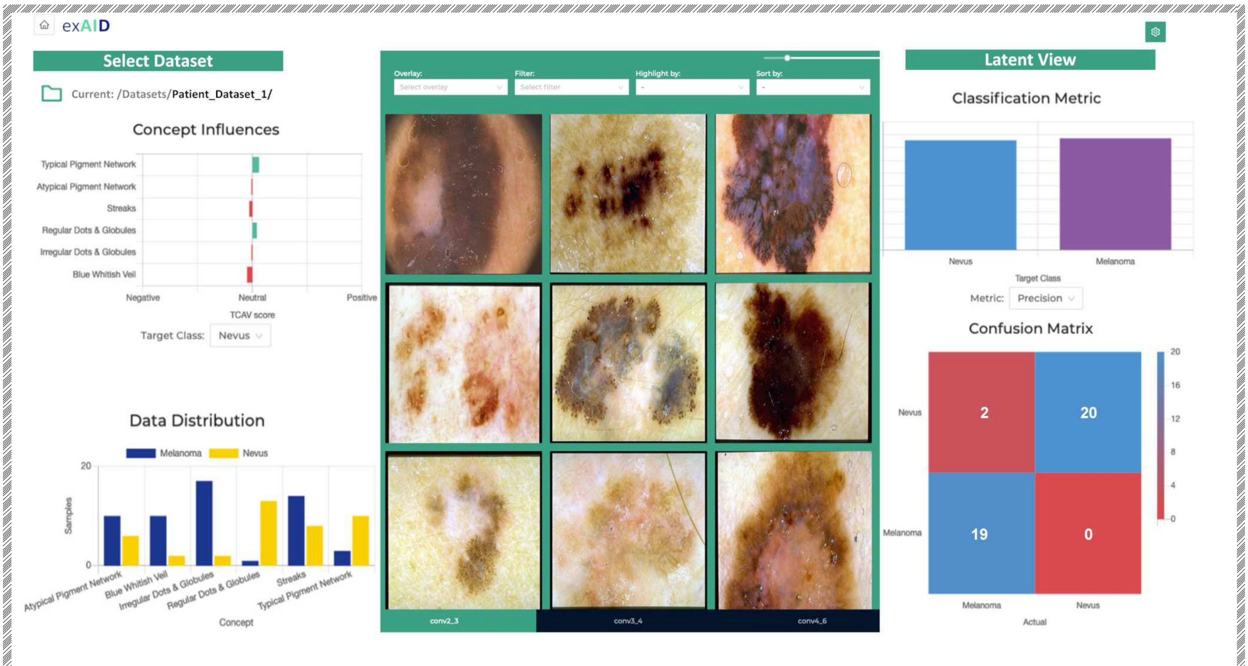


In Segment Classifier module, ConvLayer was replaced with Dense one

Diagnostic Mode of ExAID
used as Decision Support System in clinical workflow



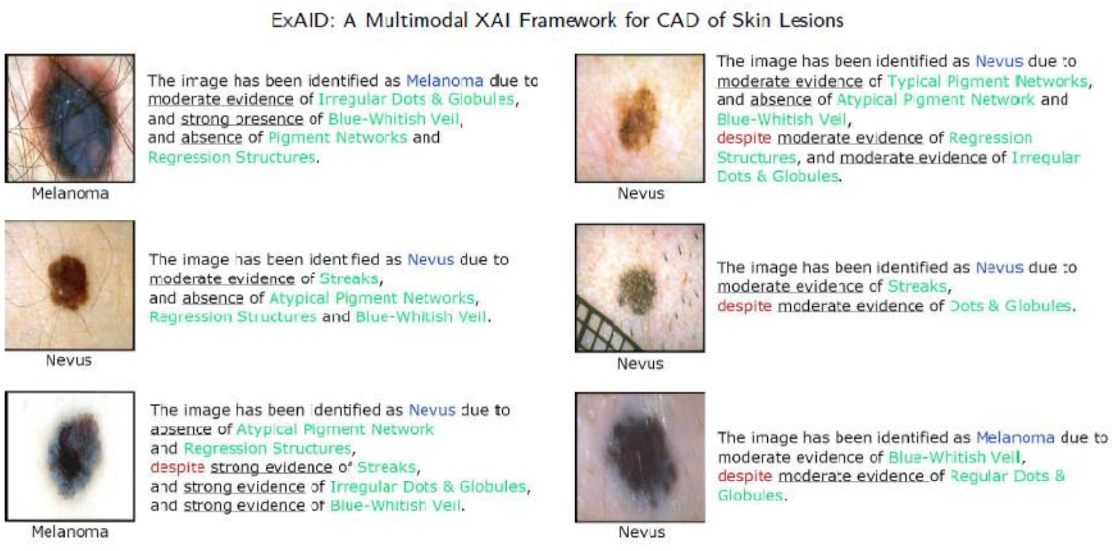
Educational Mode of ExAID



! To train resident dermatologists tohandle interactive features

xAI.Med.		2022-34

Textual explanations by ExAID
Positive and negative examples

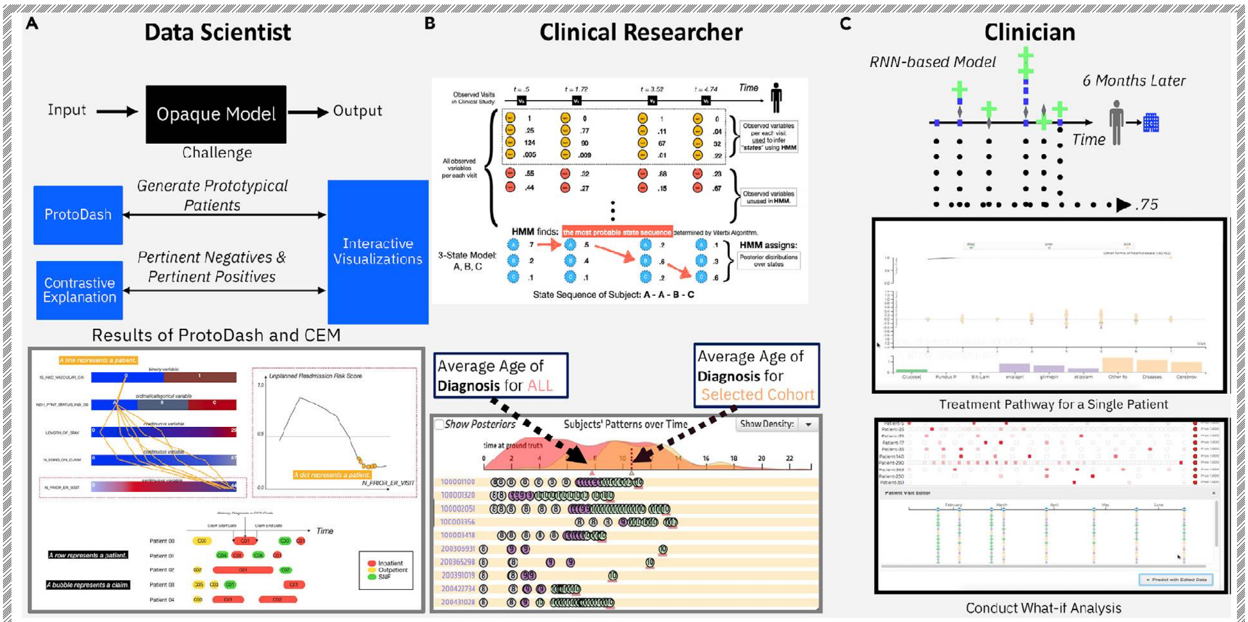


(a) Correct concept prediction. (b) Incorrect concept prediction.

- Skin lesion samples
- Ground truth class of sample is given below the image

xAI.Med .	Case Study	2022-36

Visualization and explanations



ProtoDash and CEM Explorer84 :

- ✓ Allow data scientists to inspect trained model usingcontrastive explanation method

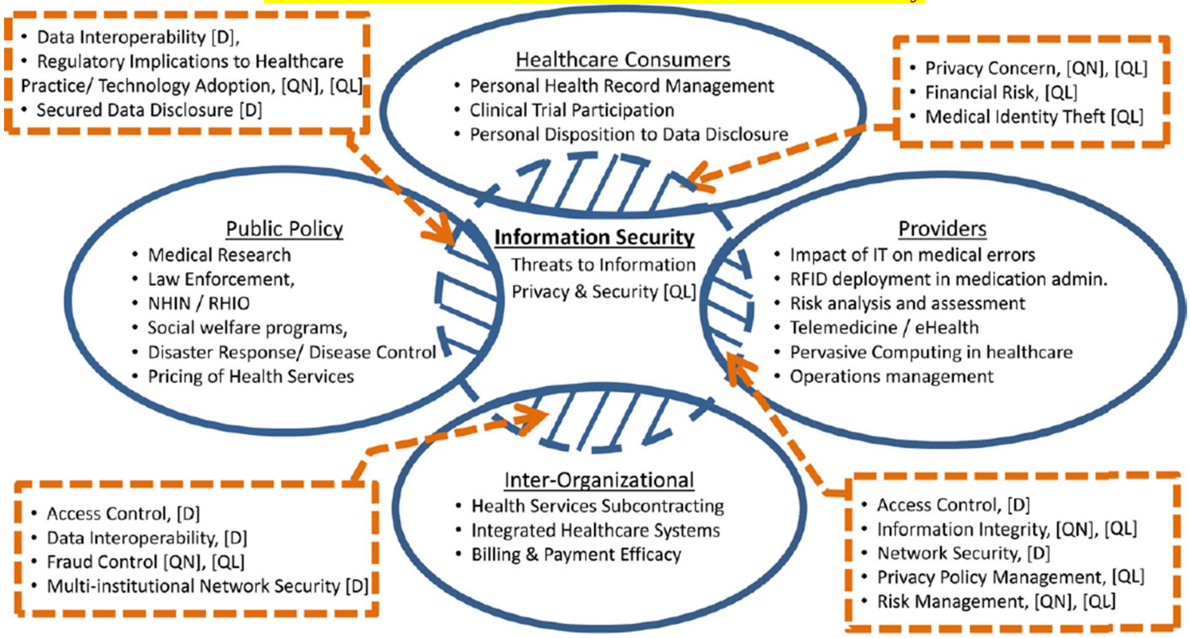
Dpvis85:

- ✓ Helps clinical researchers to understand the disease progression patterns by interacting with multiple, coordinated visualizations

Retainvis83

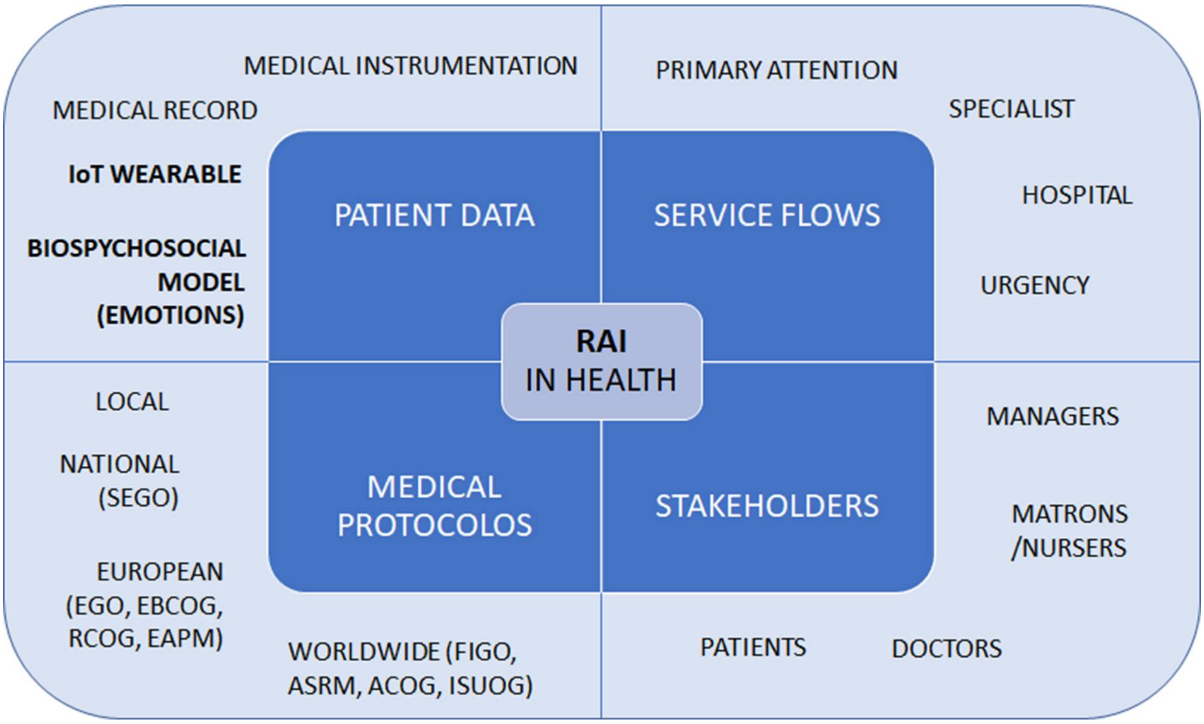
- ✓ Help clinicians test how an RNN-based model performs on a set of patients by conducting various what-if analyses

Research domains in the healthcare information security



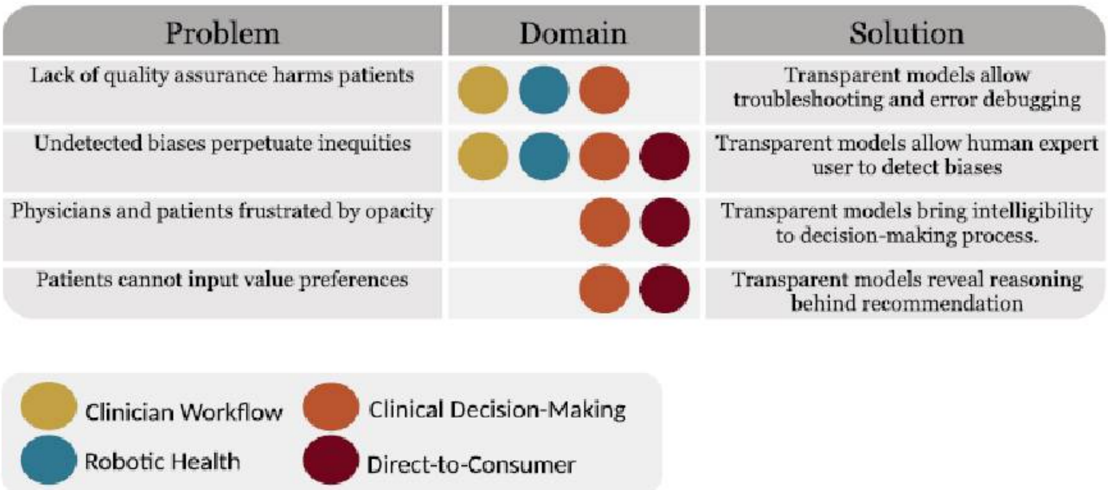
[D]: Design Research; [QL]: Qualitative Research ; [QN]: Quantitative Research

Holistic approach for Responsible Artificial Intelligence in health



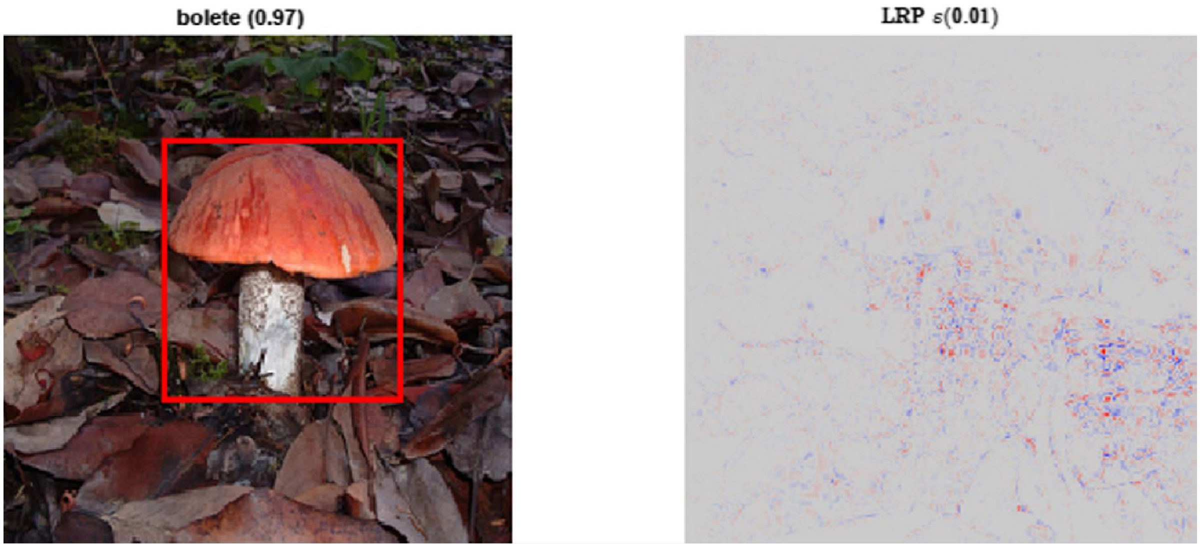
Case Study

ML—DeepLrn—DeepArch—
xAI---xxAI--
Healthcare sector



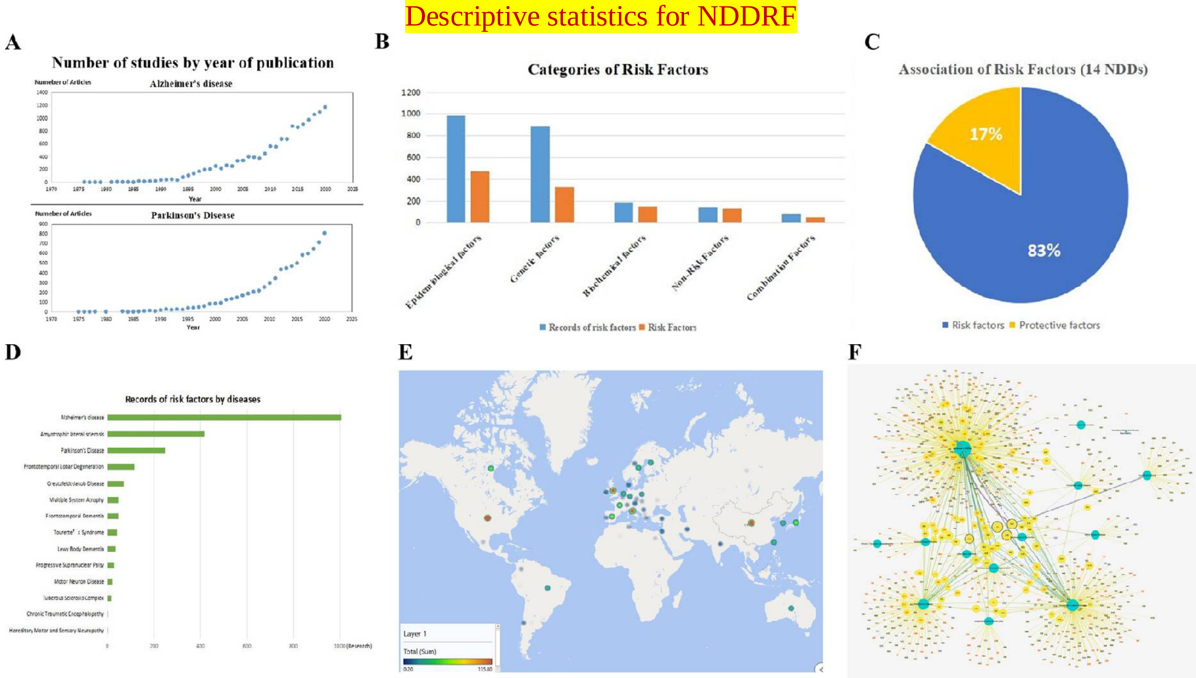
- Black-box models create problems that have the potential to undermine progress in Computational Medicine
! Remedy :eXplainable models -- x output – xAi –xAdvices-on-time (xART)
+ Transparent models offer solutions to these problems

xAI.Med. | 2022-58



xAI.Med. | 2022-69

Descriptive statistics for NDDRF



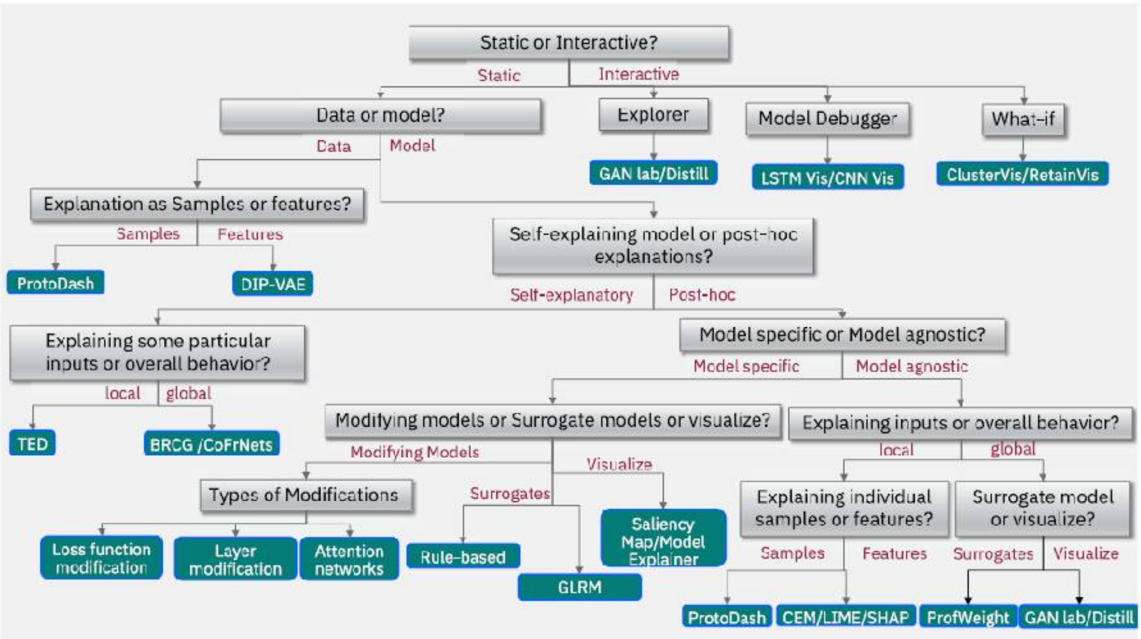
(A) # of studies by year of publication;
(B) Categories of risk factors;
(c) Association of risk factors for all 14 ndds
(d) Records of risk factors by disease in NDDRF
(E) Heat map shows country distribution
(F) Weighted network diagram -- correlation between neurodegenerative diseases and risk factors

xAI.Med **.** | 2022-36

Table 1. Summary of available open-source XAI tools

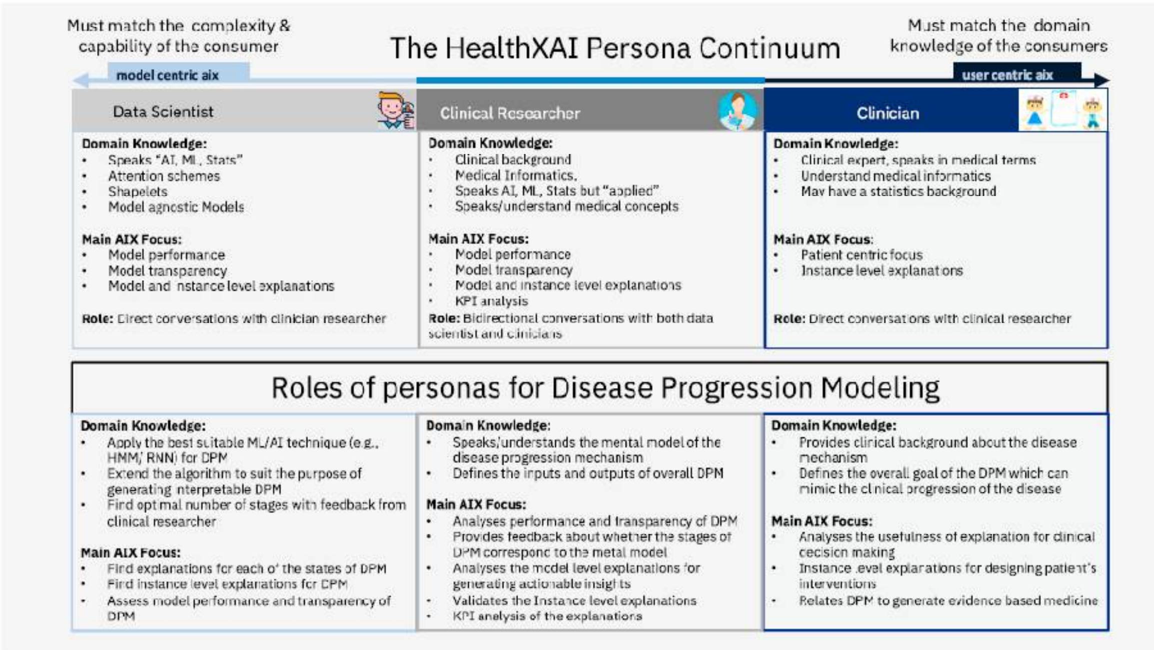
Toolkit	Data Explanations	Directly Interpretable	Self-explaining	Local *Post Hoc* Explanation	Global *Post Hoc* Explanation	Explaina- bility Metrics	URL Links
AIX 360	X	X	X	X	X	X	http://aix360.mybluemix.net
Alibi				X			https://github.com/SeldonIO/alibi
Skater		X		X	X		https://oracle.github.io/Skater/
H2O		X		X	X		https://github.com/h2oai/mli-resources
InterpretML		X		X	X		https://github.com/interpretml/interpret
EthicalML-XAI					X		https://github.com/EthicalML/xai
DALEX				X	X		https://modeloriented.github.io/DALEX/
tf-explain				X	X		https://github.com/sicara/tf-explain
iNNvestigate				X			https://github.com/albermax/innvestigate
modelStudio	X	X		X	X		https://bit.ly/3uOnU5y
ELI5		X		X	X		https://github.com/TeamHG-Memex/eli5
Iml		X		X	X		https://bit.ly/3iBv8Vx

Taxonomy tree for explainability in AI models



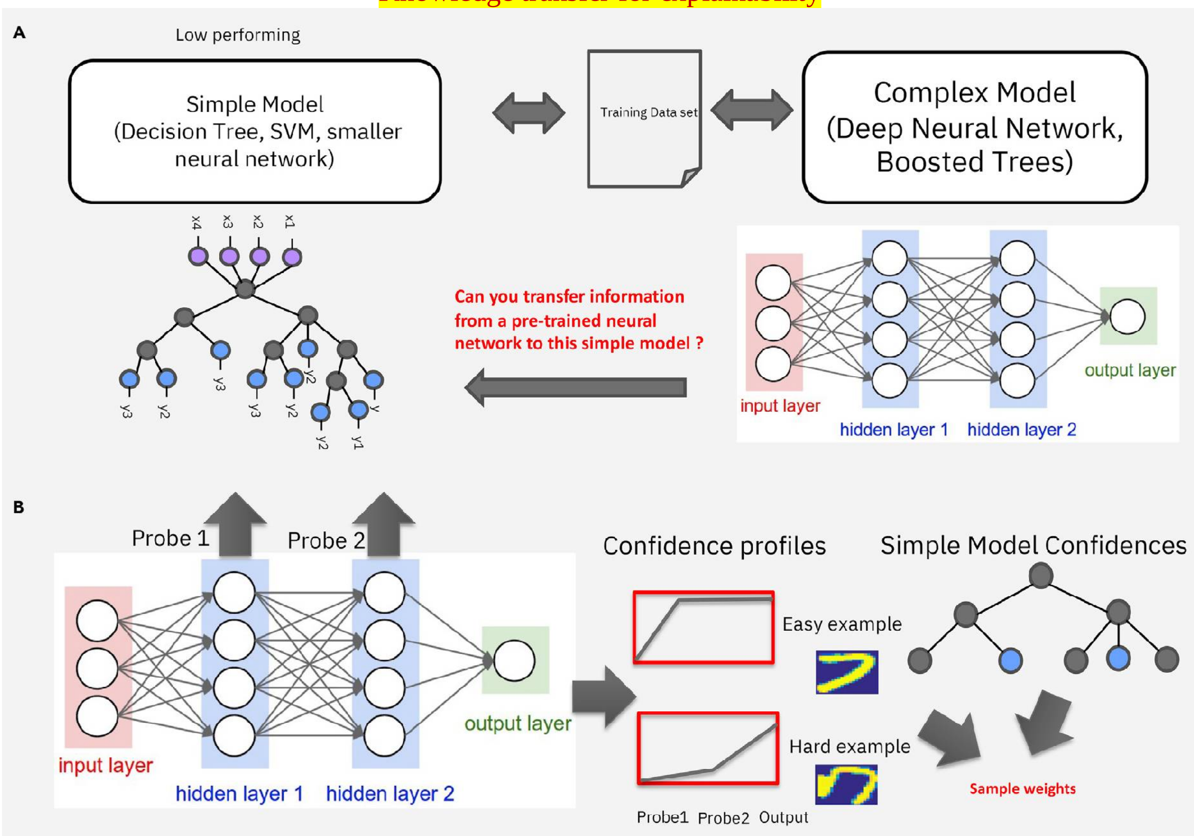
	Rule-Base for choice of Explanation methods
If	Goal is to explain models instead of data
Then	Whether a local explanation for individual samples OR Global explanation for the entire model

If	Global path
Then	Should it be a post hoc method or a self-explaining one
If	Self-explaining method
Then	TED (teaching explanations for decision making)
If	global method
Then	BRCG (Boolean rule sets with column generation)
If	Model agnostic.
Then	Post hoc
	If so, | ? Explaining in terms of samples or features
If	Sample side
Then	Which prototypes
If	Feature methods
Then	Choices among the contrast of explanations methods (CEMs) OR LIME37 or SHAP
If	Post hoc global methods
Then	surrogate models, such as ProfWeight, are available.
If	Model-specific methods
Then	Modifying models, OR Surrogate models, OR Simply visualizations
If	Global explanations forentire model
Then	Posthoc OR Directly interpretable model?
If	Directly interpretable model
Then	Boolean rule set, such as BRCG OR GLRMs (generalized linear rulemodels))



xAI.Med. | 2022-36

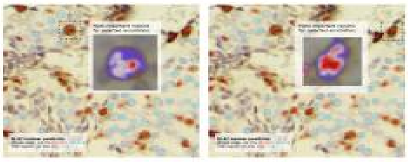
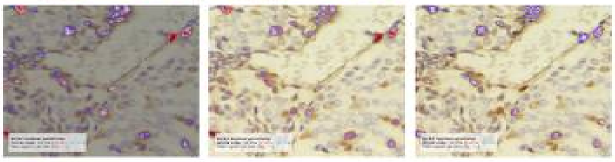
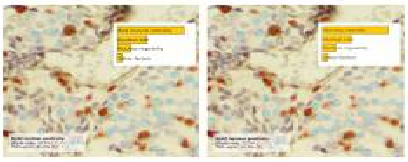
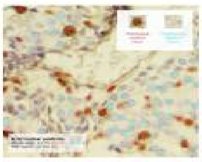
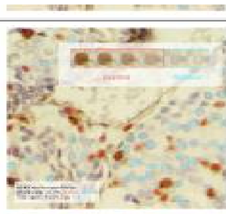
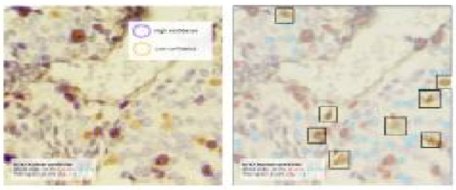
Knowledge transfer for explainability



xAI.Med. | 2022-41

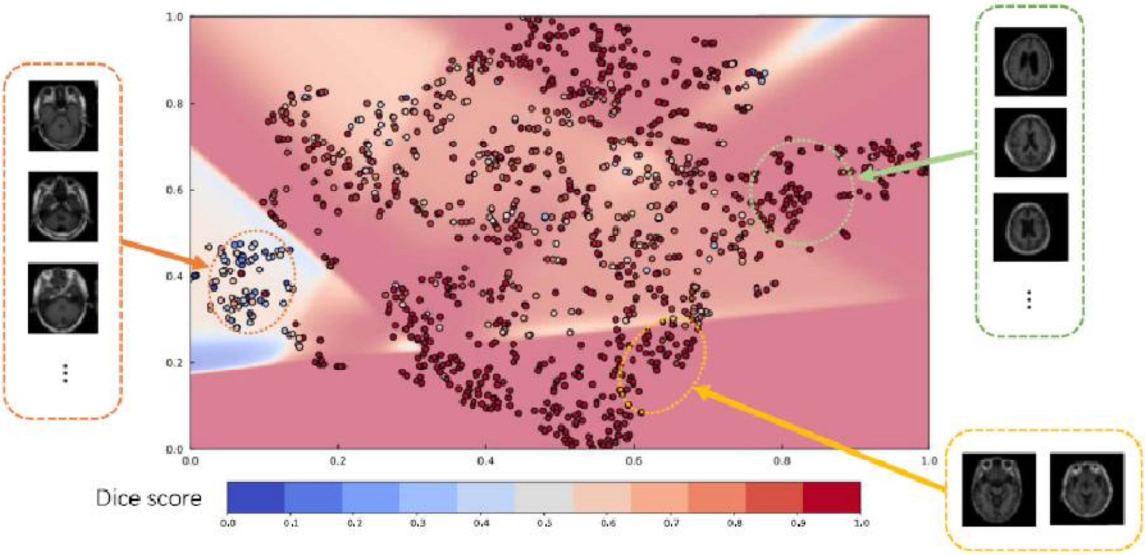
Explanation instances

Xplain_style: [plain-text ; image variants]

Name	Description	Variants
Saliency Map (Local)	Show the most relevant pixels for the classification of a selected annotation	
Saliency Map (Global)	Show the most relevant pixels for the positive classifications within this region of interest	
Concept Attribution	Show the most important features attributed to positive classifications	
Prototypes	Show prototypical positively and negatively classified annotations within this region	
Counterfactuals (One-axis)	Show generated examples interpolating between positive and negative examples, showing model classifications for each	
Counterfactuals (Two-axis)	Show generated examples changing in two principal factors of variation, showing model classifications for each	
Trust Scores	Display low-confidence annotations for review	

Dice scores

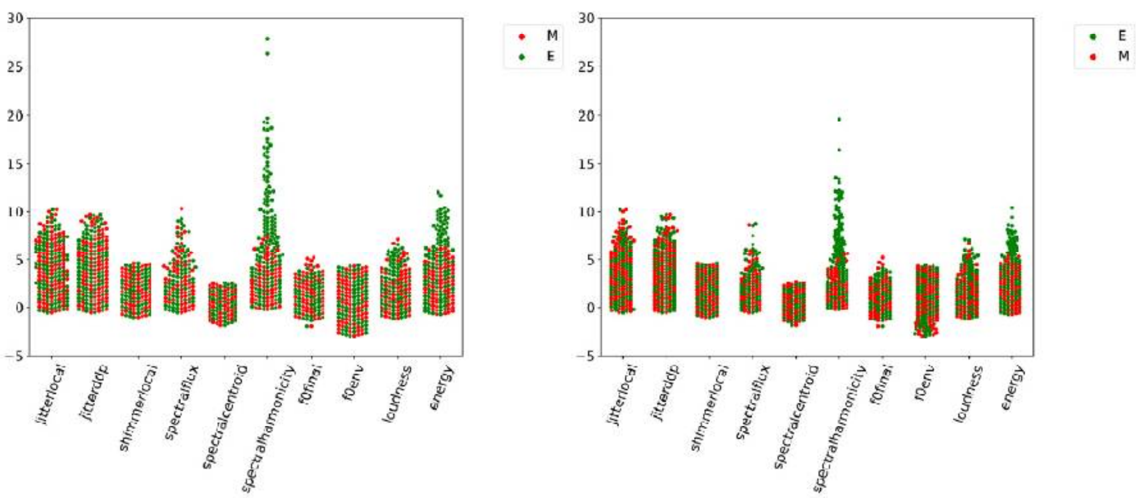
xAI.Med.	2022-62



✓ Samples from different regions of the plane are presented on the left-hand side and right-hand side

Swarm plot

Swarmplot visualization

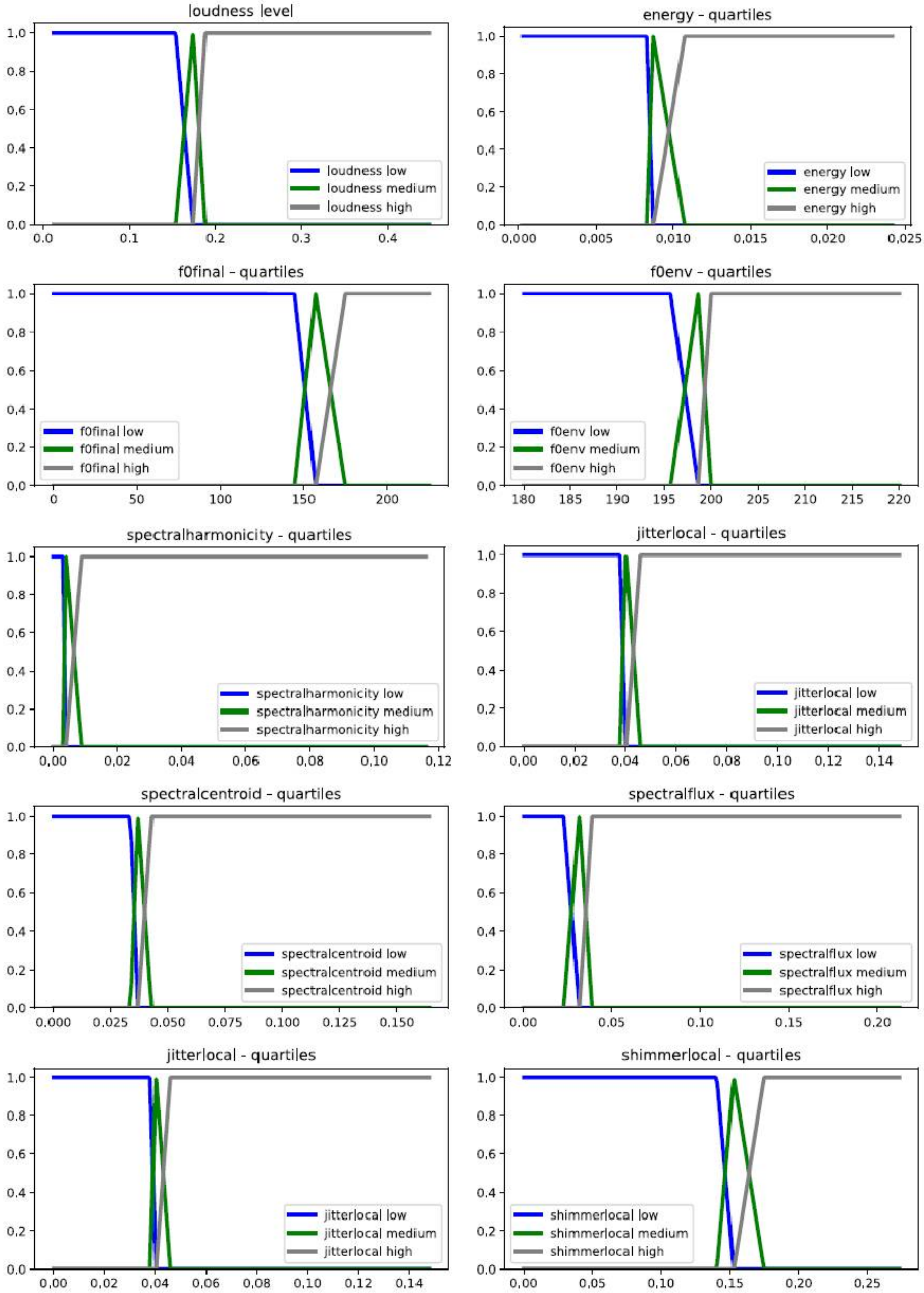


(a) cluster 0 (b) cluster 1

fuzzy Plot

Linguistic variables

Linguistic variables and their fuzzy sets derived from prototypes

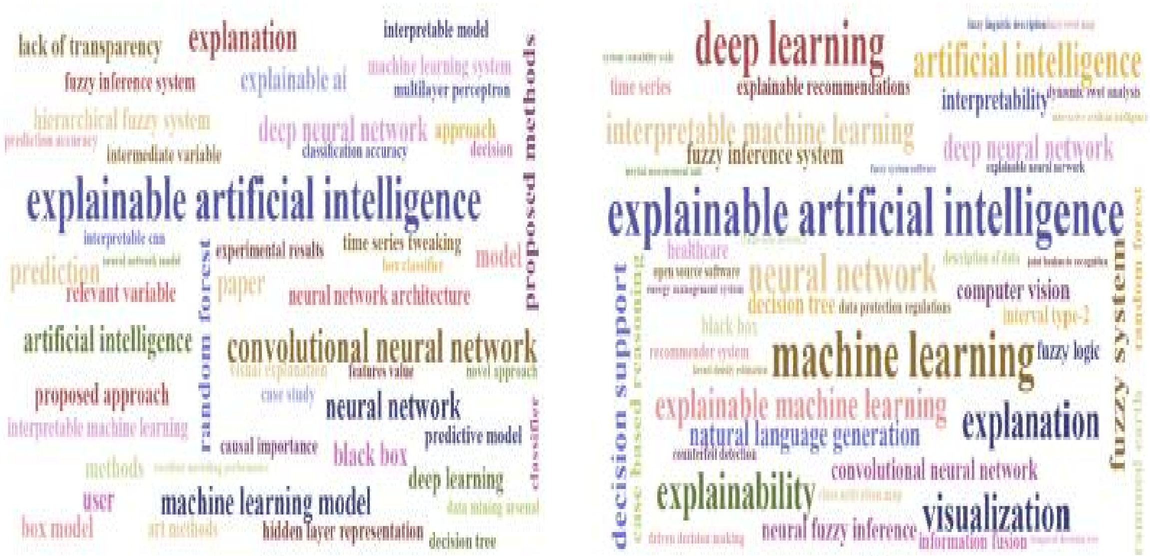


Systematic Literature Search -- Advances

xAI.Med.		2022-204

Word cloud

Font size is proportional to the number of occurrences of the terms
different colours are used to discriminate terms with equal font size

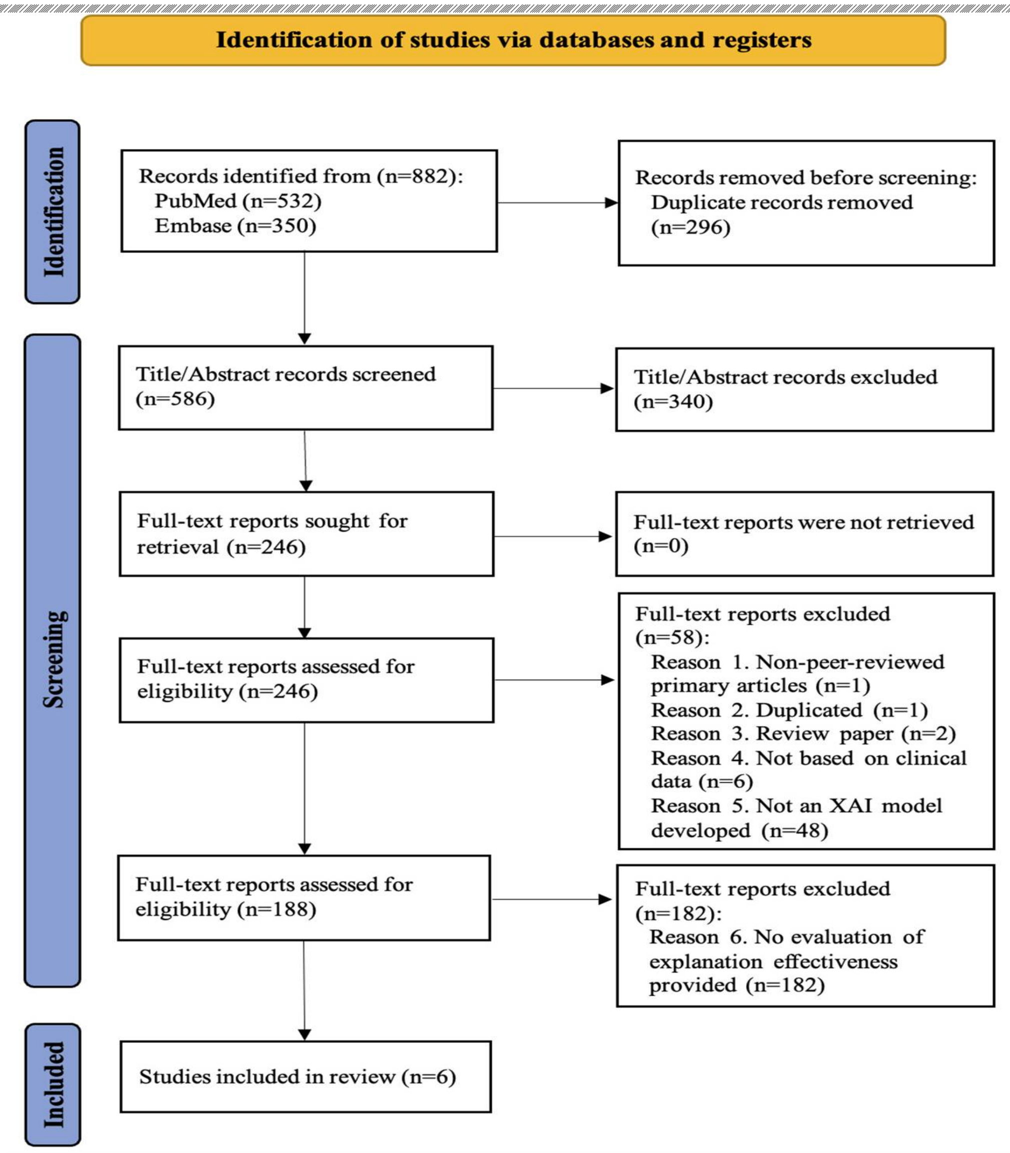


(a) (b)

(a) author-defined keywords	b) keywords extracted from abstracts through NLP

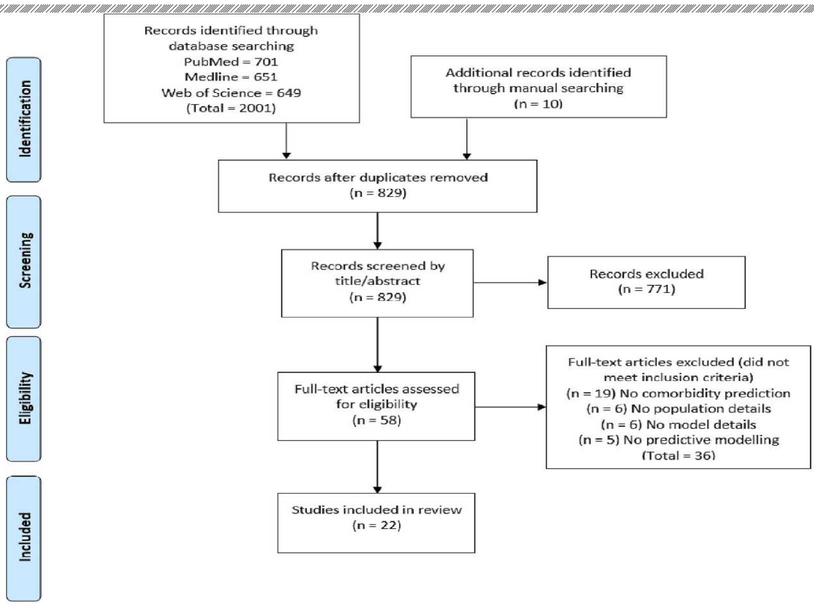
- Terms from keywords are more conceptual
- Abstracts contained specific terms on the methods and tasks
- Summarily, illustrate remarkable number of terms of XAI

xAI.Med.		PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis)	2023-28
		PRISMA flow for identification, screening, and inclusion of literature	



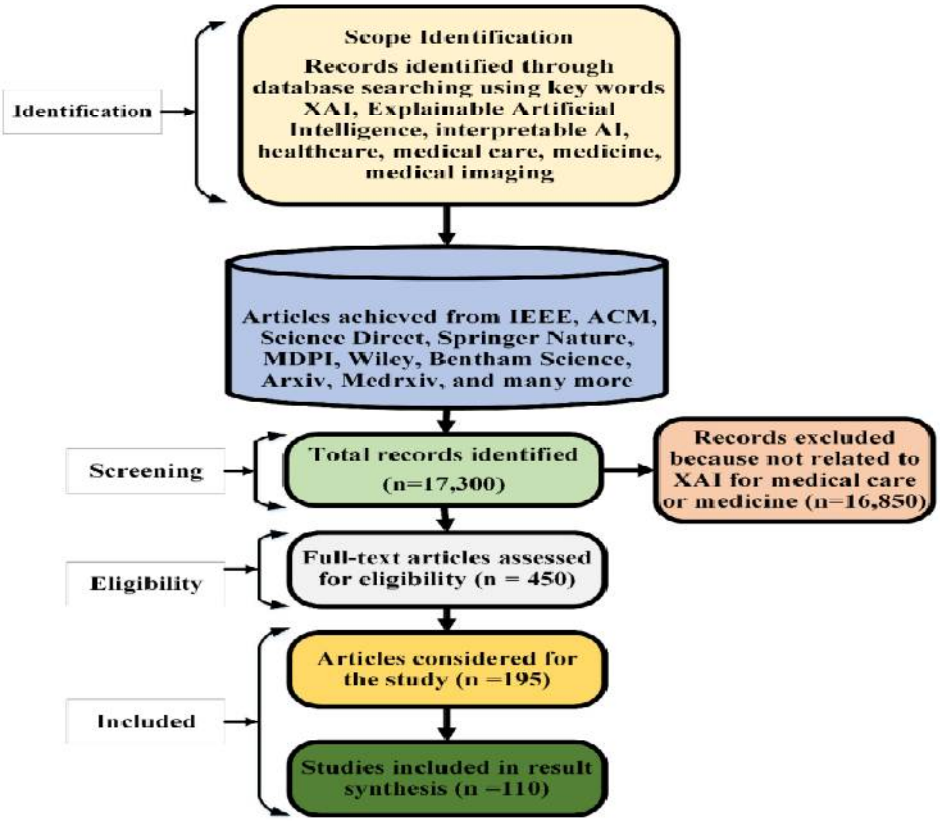
xAI.Med.	2023-19

PRISMA flow diagram for primary research journal publications



xAI.Med.		2023-101

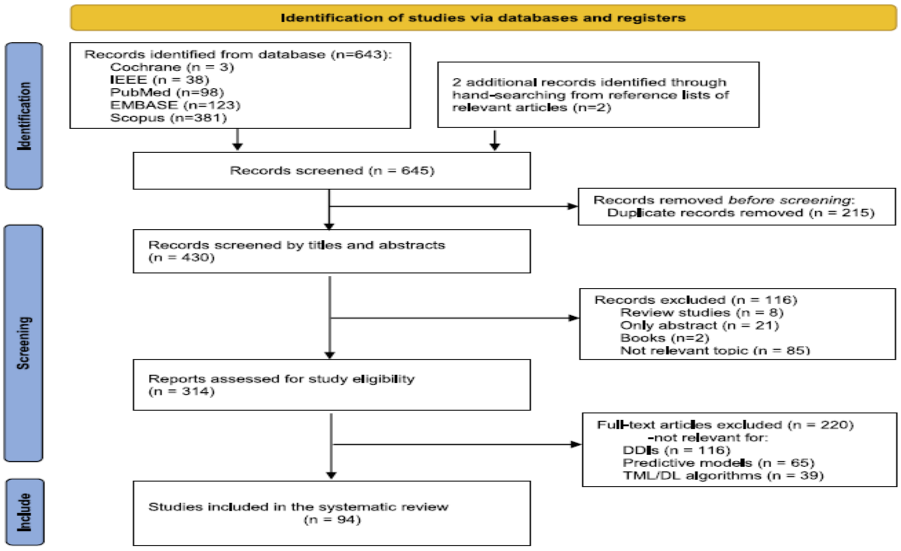
PRISMA guideline for manuscript selection



PRISMA : preferred reporting items for systematic reviews and meta-analyses

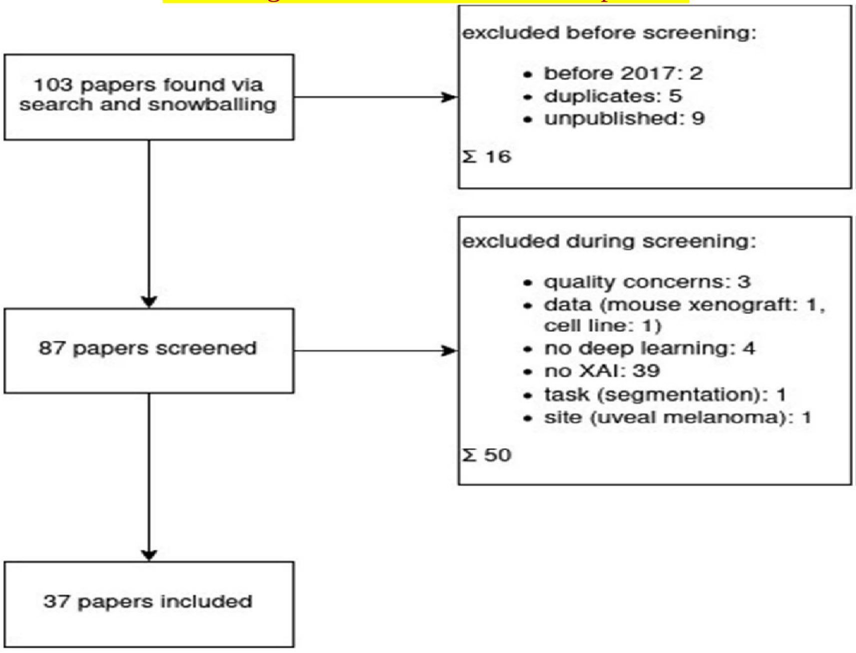
xAI.Med.		2022-39

PRISMA diagram showing our literature strategy search



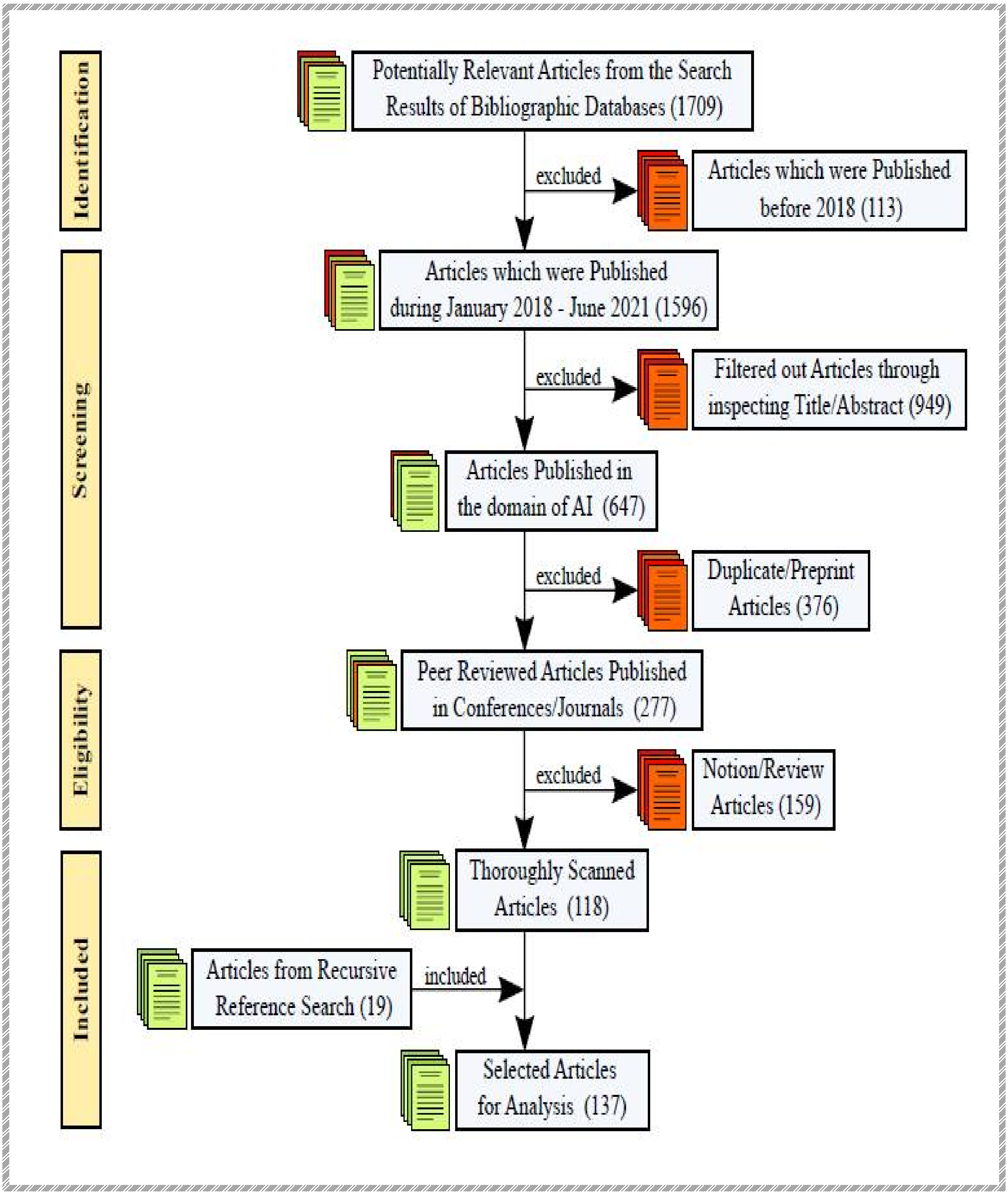
xAI.Med.		2022-51

Flow diagram of the article selection process



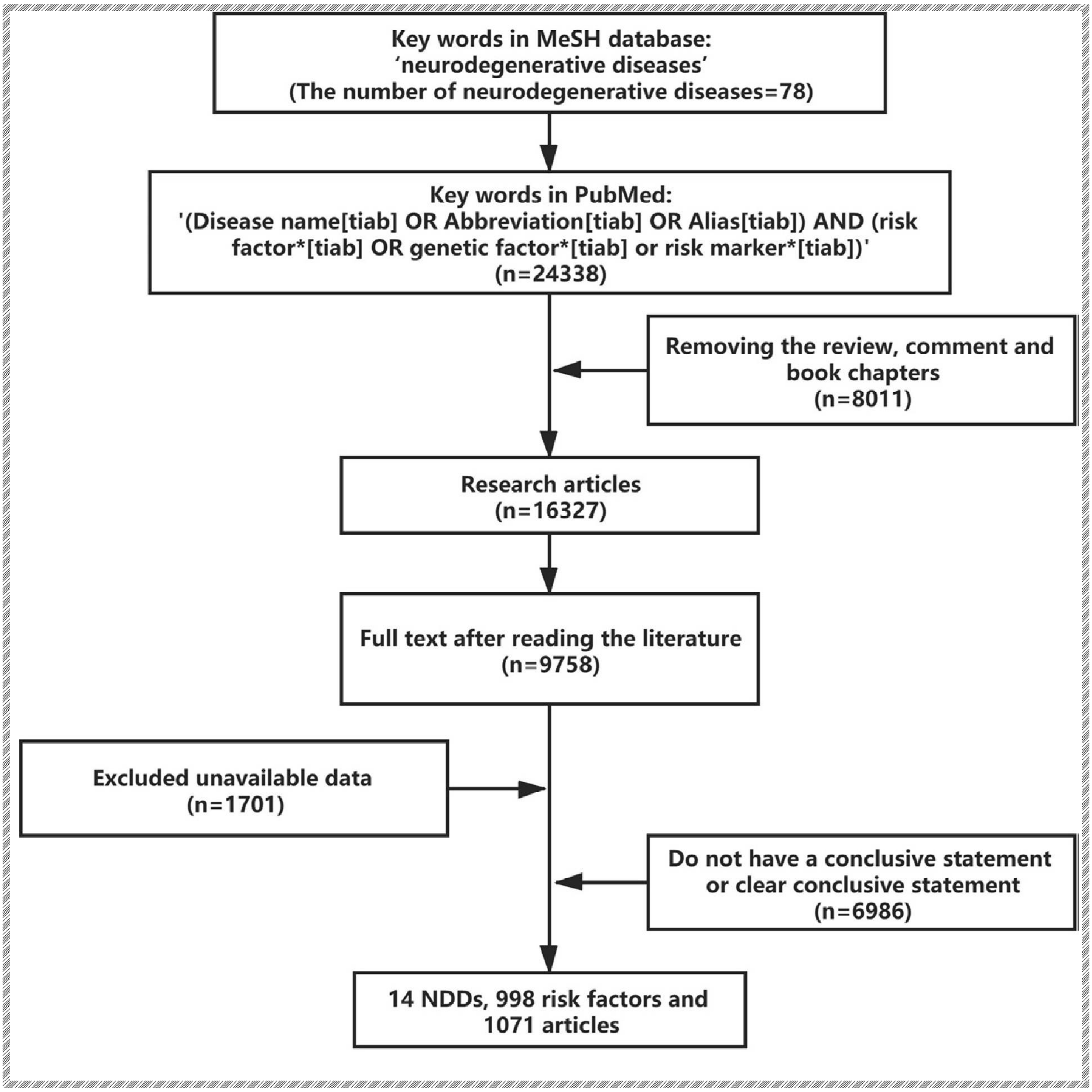
xAI.Med.		2022-204

Research article selection by PRISMA protocol



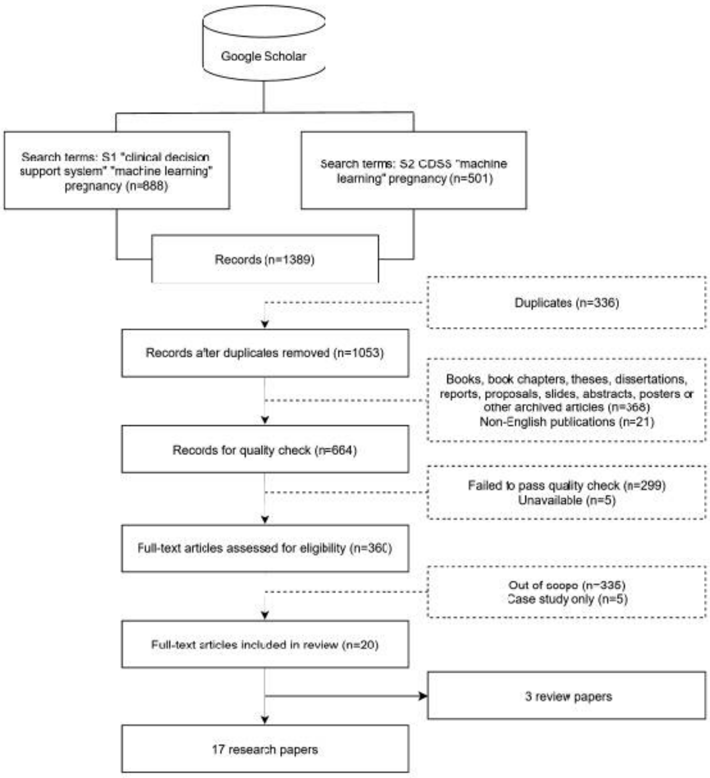
xAI.Med .		2022-69

Flowchart for collection of risk factors for neurodegenerative diseases



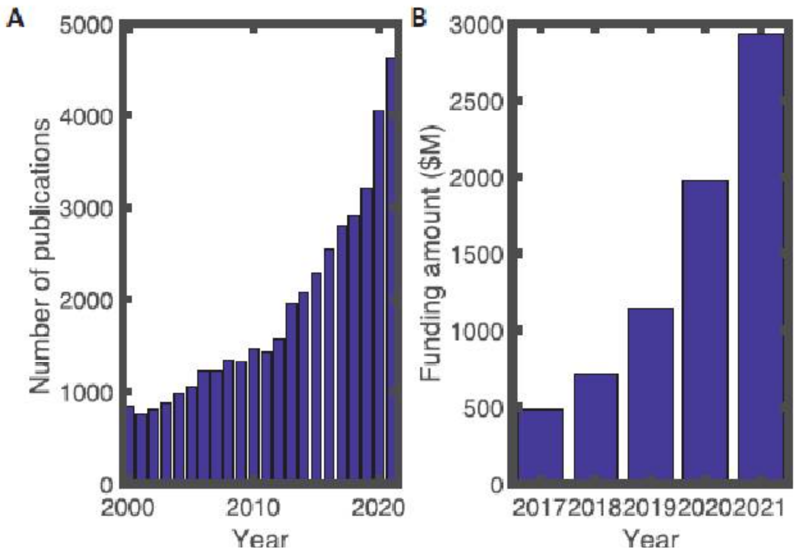
xAI.Med.		2023-01

Flow diagram summarising the process of conducting searches, paper selection and filtering



xAI.Med. | 2022-10

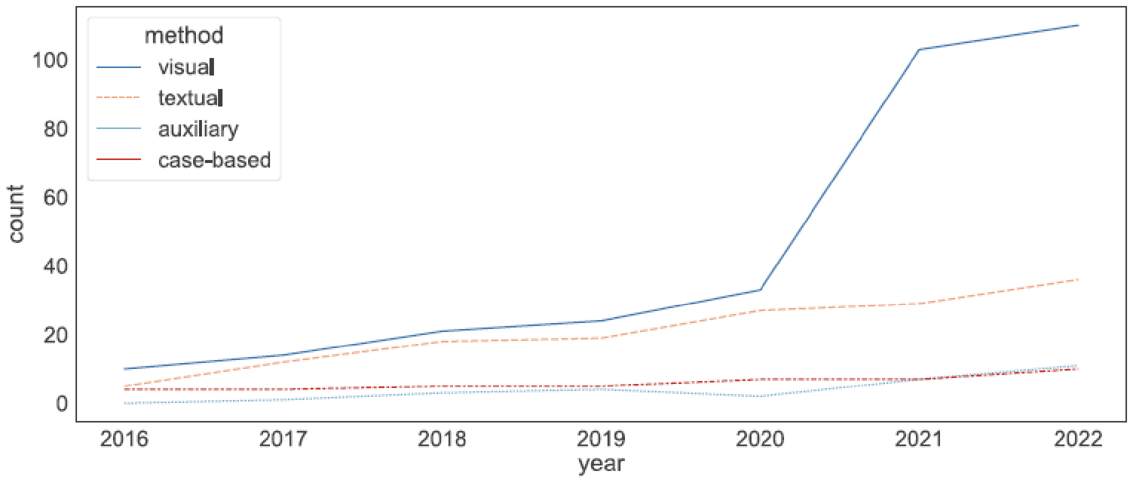
ML research in mental health and categorization of neuroimaging



- ✓ The number of PubMed publications with keywords "machine learning or AI" and "psychiatry or mental health" in the title or abstract (years 2000–2021).
- ✓ Growth of mental health tech funding in the US market (years 2017–2021)
  - data source: https://www.cbinsights.com

xAI.Med. | 2023-22

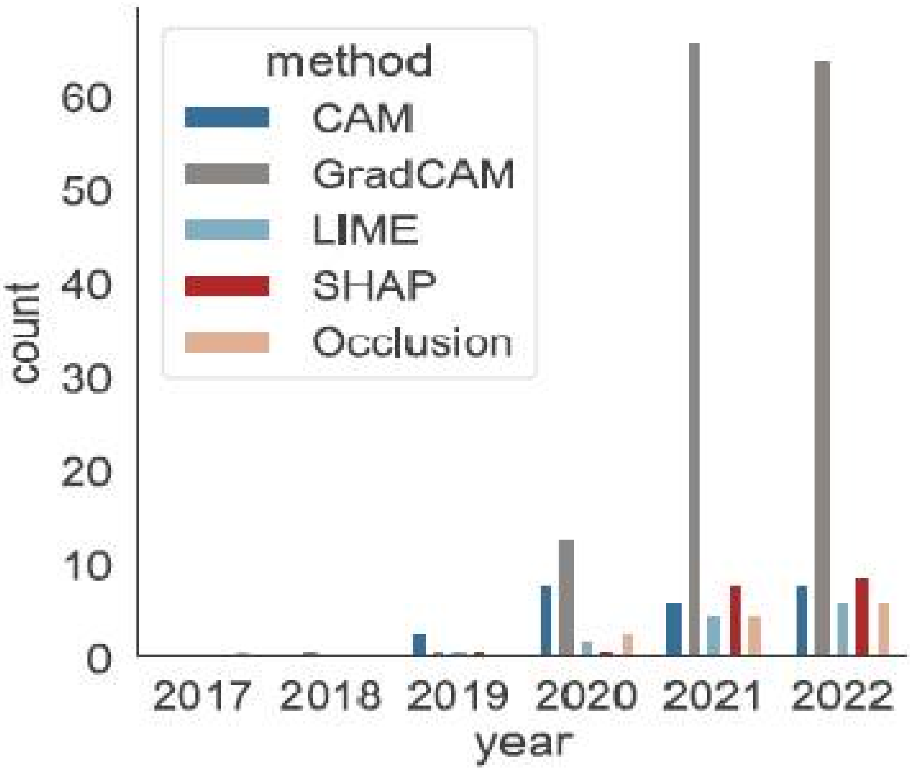
Literature trend for Medical images



✓ Year wise reports of saliency-based (visual) and other non-visual (textual, example-based-case-based) XAI methods

xAI.Med. | 2023-22

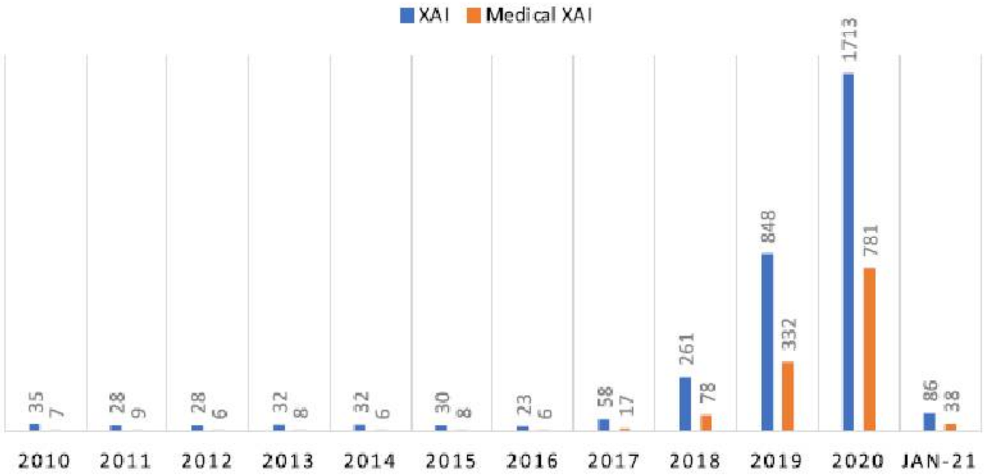
Saliency-based XAI methods

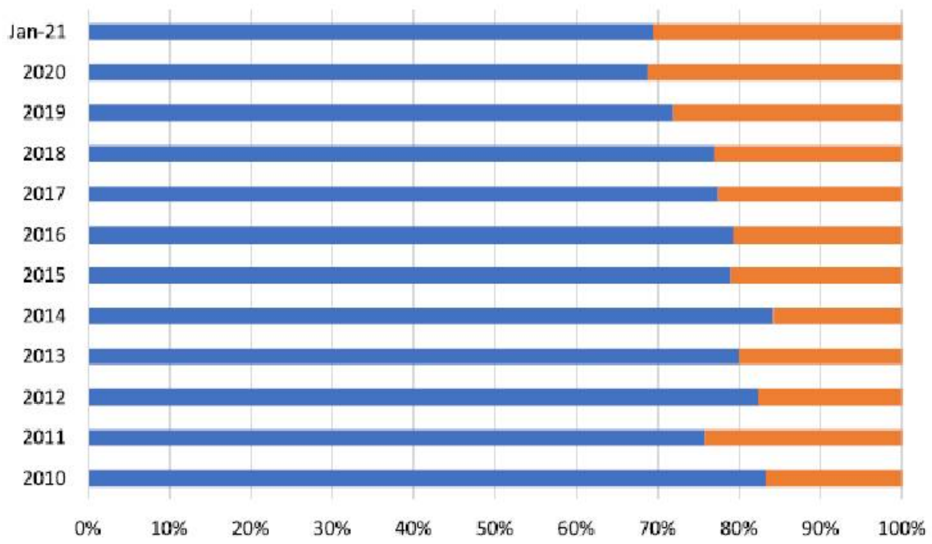


Medical image analysis

xAI.Med. | 2022-

Publication per year for XAI and medical XAI (top)
percentage for two categories of research (bottom)





Data retrieved from Scopus® (Jan 8th, 2021)

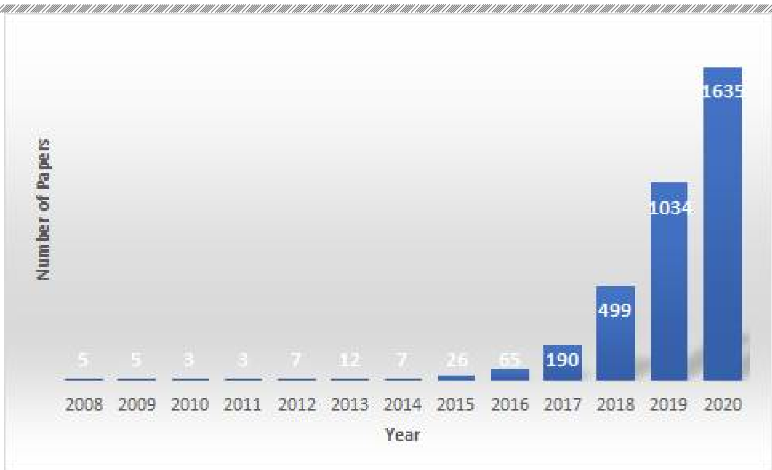
commands when querying this database Scopus®—XAI:

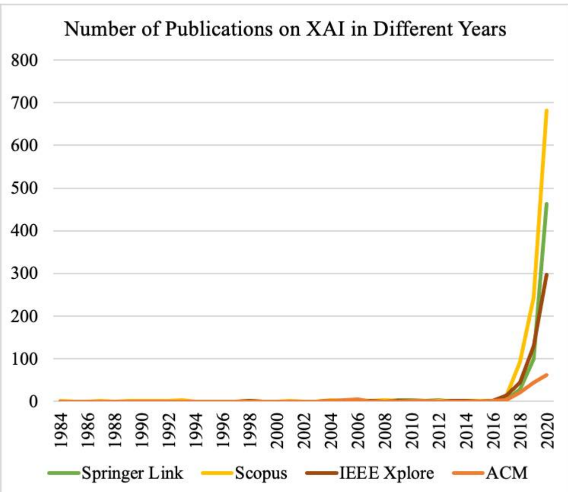
- ✓ (ALL(‘‘Explainable AI’’) OR ALL(‘‘Interpretable AI’’) OR ALL(‘‘Explainable Artificial Intelligence’’) OR ALL(‘‘Interpretable Artificial Intelligence’’) OR ALL(‘‘XAI’’)) AND PUBYEAR = 20XX;

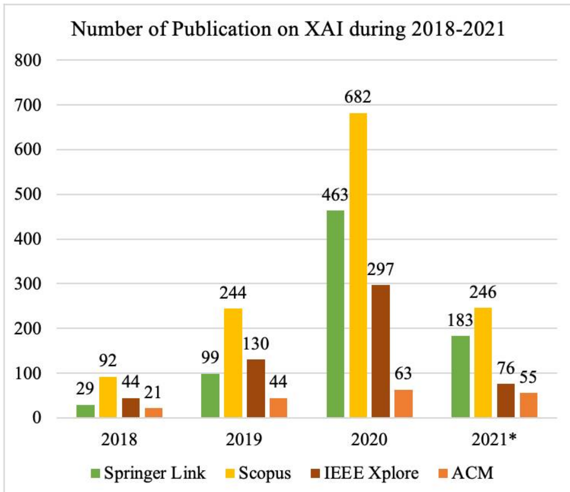
- ✓ Medical XAI: (ALL(‘‘Explainable AI’’) OR ALL(‘‘Interpretable AI’’) OR ALL(‘‘Explainable Artificial Intelligence’’) OR ALL(‘‘Interpretable Artificial Intelligence’’) OR ALL(‘‘XAI’’)) AND (ALL(‘‘medical’’) OR ALL(‘‘medicine’’)) AND PUBYEAR = 20XX, in which XX represents the actual year

xAI.Med.		2022-208

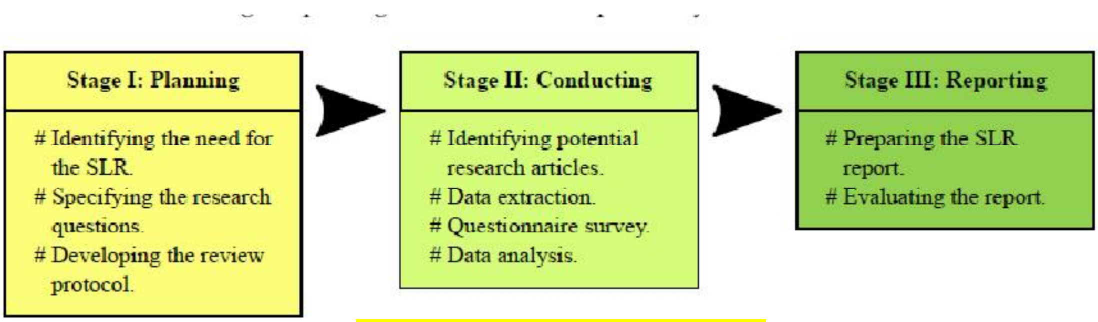
Published papers using deep learning models
to predict cancer risk, classification, prediction and
detection per year since 2008
from PubMed





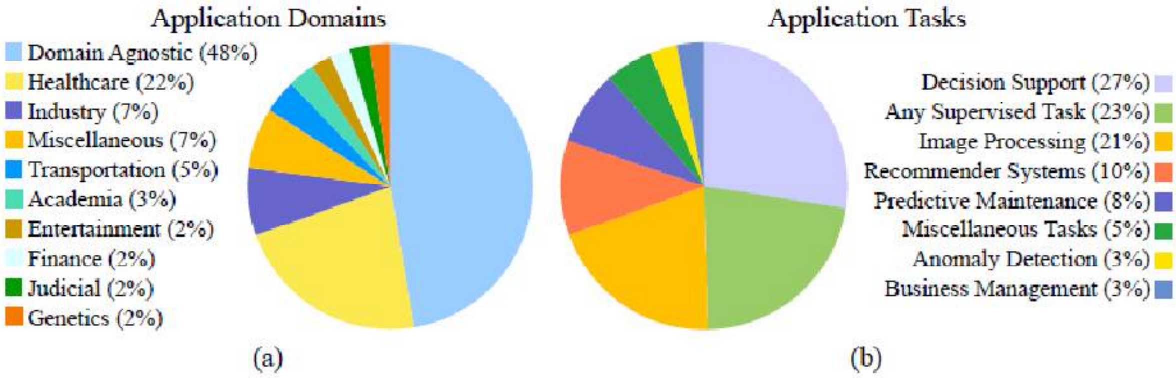


- ✓ y axis : Number of published articles on XAI
- ✓ x axis : four bibliographic databases in recent decades



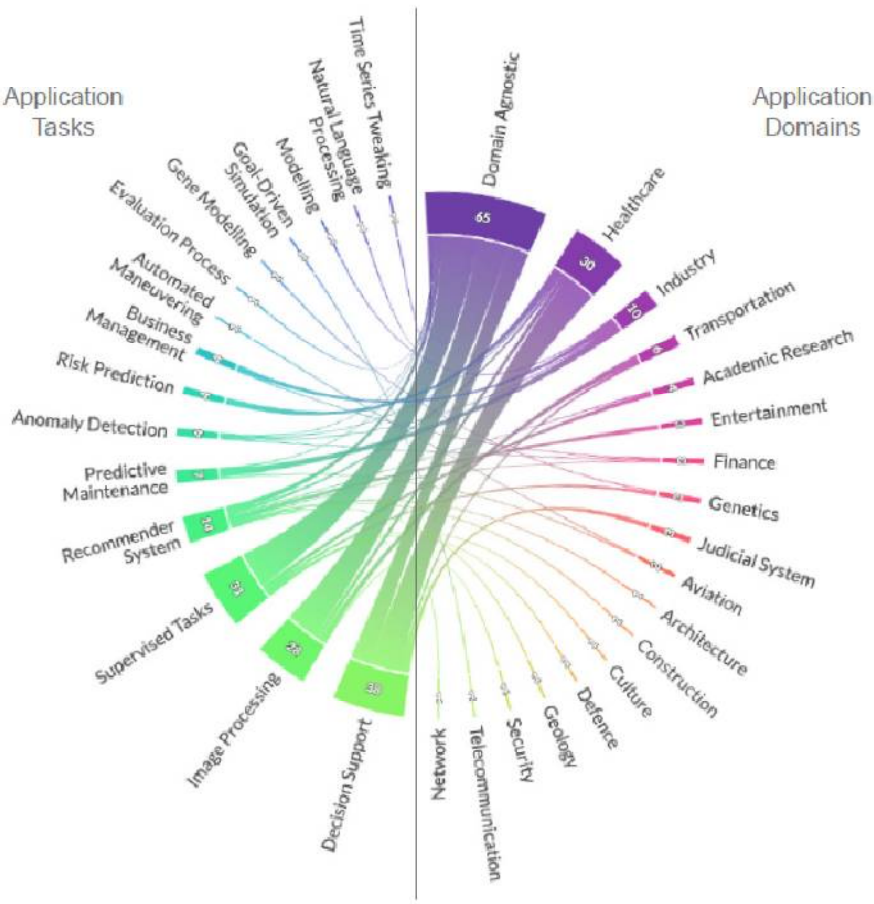
Systematic literature review (SLR)

Percentage of the selected articles on different XAI methods for different application



xAI.Med.		2022-204

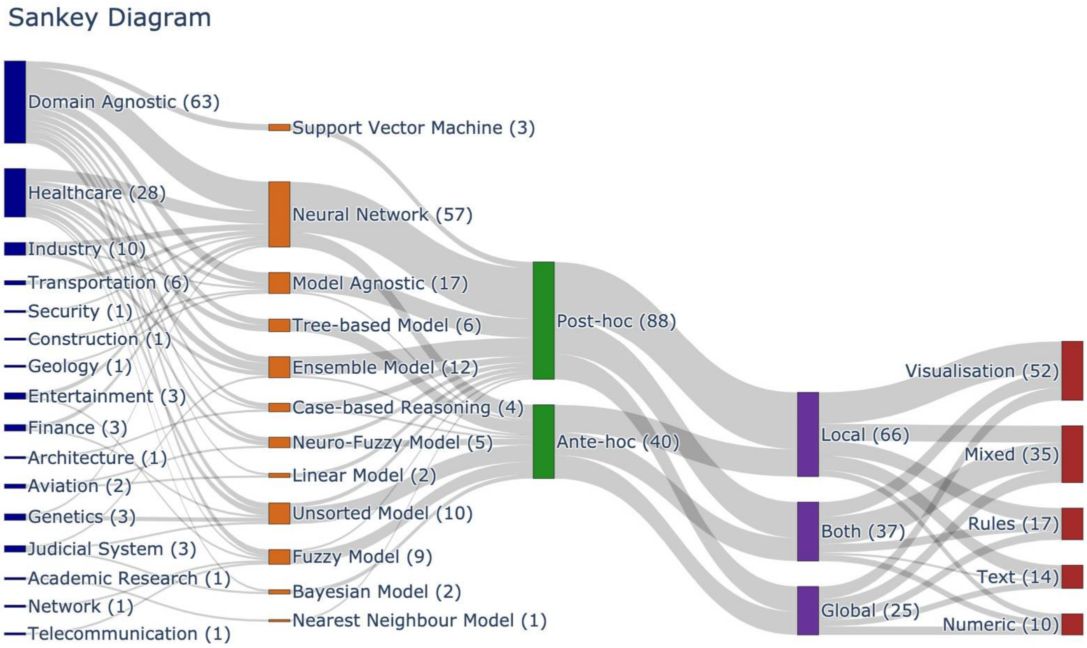
Chord diagram



Number of selected articles published on the XAI methods and evaluation metrics from different application domains

xAI.Med.		2022-204

Application domains and clustered on the basis of
AI/ML model type, stage, scope, and form of explanations

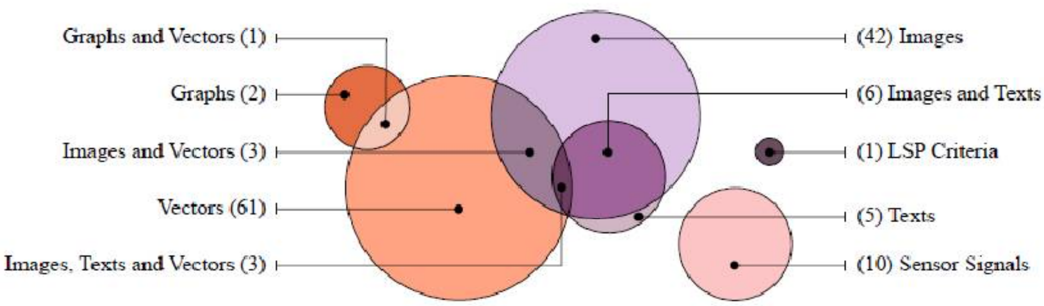


✓ Number of articles with each of the properties is given in parentheses

xAI.Med.		2022-2004

XAI methodologies

.Venn diagram with the number of articles using different forms of data



✓ The sizes of the circles are approximately proportional to the number of articles (shown within parentheses)

xAI.Med.		2023-28

Explanation types

Authors, Year	Research aims	Data resources	Data types	Input	Output	AI methods	AI Performance metrics	Explainable techniques
Alahmadi, A. et al., 2021 [29]	To develop an explainable rule-based decision tree classification model to automate the detection of QT-prolongation at risk of Torsades de Pointes (TdP)	Public dataset, clinical trial approved by Food and Drug Administration (FDA) in 2014	ECG image data	ECG	Classification of Torsade de Pointes (TdP)	Rule-based algorithm	Accuracy Balance Sensitivity Specificity PPV F1-score ROC (AUC) Precision-Recall (AUC) MCC Error rate	Pseudo-coloring methodology
Born, J. et al., 2021 [30]	To develop an explainable classification model for differential COVID-19 diagnosis	Public dataset, Lung Point-Of-Care Ultrasound (POCUS)	Ultrasound video data	Ultrasound	Classification of COVID-19	CNN	Precision Recall F1-score Specificity MCC	CAM
Neves, I. et al., 2021 [31]	To develop an explainable ECG classification model on time series	Public dataset, Massachusetts Institute of Technology-Beth Israel Hospital (MIT-BIH) Arrhythmia	ECG image data	ECG	Classification of arrhythmia	KNN CNN	F1-score Precision Recall AUC	PFI LIME SHAP
Sabol, P. et al., 2020 [32]	To develop an explainable classification model for colorectal cancer diagnosis	Public dataset, Colorectal cancer pathology image	Histopatho-logical image data	Colorectal cancer pathology image data	Classification of colorectal cancer	CNN	Accuracy Precision Recall F1-score	CFCMC
Tan, W. et al., 2021 [33]	To develop an explainable deep learning model for the automatic diagnosis of fenestral OS	EHR data, the Fudan University	CT scan image data	Temporal bone high-resolution computed tomography (HRCT)	Classification of fenestral otosclerosis	Conventional image processing algorithm	Accuracy Sensitivity Specificity PPV NPV	Faster-RCNN
Derathé, A. et al., 2021 [34]	To explain the previously developed prediction model for surgical practice quality	EHR data, the CHU Grenoble Alpes Hospital	Laparoscopic sleeve gastrectomy (LSG) operation video data	Laparoscopic operation videos	Extraction of the most important variables to predict the quality of surgical practice	SVM	Accuracy Sensitivity Specificity	Value-permutation and Feature-object semantics

Definitions of explainable techniques

Explainable techniques	Mechanism
Permutation Feature Importance (PFI)	The PFI is a technique for overall interpretability by examining the model score after shuffling a single feature value [31].
Local Interpretable Model-agnostic Explanation (LIME)	The LIME is a perturbation-based strategy that uses a surrogate interpretable model to substitute the complex model locally, providing local interpretability [31].
SHappley Additive exPlanation (SHAP)	The SHAP is a method for determining how each feature contributes to a specific outcome [31].
Faster Region with Convolutional Neural Network (R-CNN)	The faster R-CNN presented the Region Proposal Network [RPN], which speeds up the selective search. RPN adheres to the last convolution layer of CNN. Proposals from RPN are given to a region of interest pooling (RoI pooling), then classification and bounding-box regression [56].
Pseudo coloring methodology	The pseudo coloring methodology employs a range of colors to represent continuously changing values [29].
Class Activation Map (CAM)	The CAM uses global average pooling to generate class-specific heatmaps that indicate discriminative regions [30]
Value permutation and Feature-object semantics	The permutation of values is analyzed for their impact on predictions, and the most significant variables are then translated into statements using feature-object semantics [34].
Cumulative Fuzzy Class Membership Criterion (CFCMC)	The CFCMC offers a confidence measure for a test image's classification, followed by a representation of the training image and the most similar images [32].

Database	Query	Results
PubMed	"explainable artificial intelligence" [Title/Abstract] OR "xai" [Title/Abstract] OR "explainable ai" [Title/Abstract] OR "interpretable ai" [Title/Abstract] OR "interpretable artificial intelligence" [Title/Abstract]	532
Embase	'explainable artificial intelligence':ab,ti OR xai:ab,ti OR 'explainable ai':ab,ti OR 'interpretable ai':ab,ti OR 'interpretable artificial intelligence':ab,ti	350

xAI.Med. | 2022-204

Stage, scope, and form of explanations