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

ISSN: 2278-1862

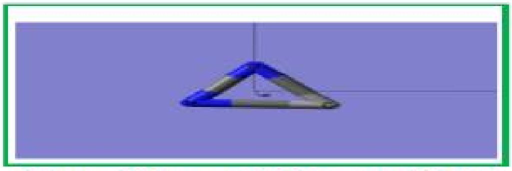
Journal of Applicable Chemistry

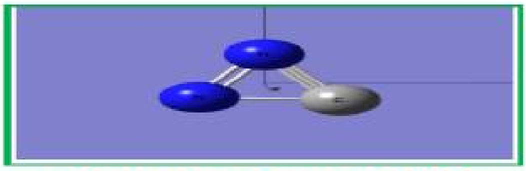
2023, 12 (4): 500-552

(International Peer Reviewed Journal)

New Chemistry News

⁻N=C=N⁻





New News of Chem (NNC) **ChemNewsNew (CNN)**

CNN-55--Fit (Figure Image TableScript...)Bases (Bfit) Part 3.xAI. Architectures . Medicine – part-1

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. SomasekharaRao, 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: The evolution of scientific Artificial Intelligence (AI) (save consciousness) implementable on computers since mid-nineteen fifties and emergence of eXplainable AI (xAI) during last one decade is briefly described. The extensive application in most of critical research disciplines brought this approach to the fore-front-of state-of-knowledge projects including medicine, defence, industry, commerce etc. The recent trend setting xAI-probes resulted in laying confidence in AI-embedded

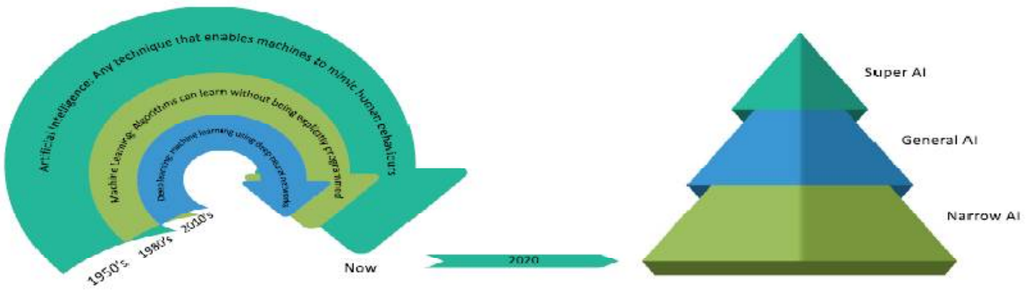
methods/products/materials. The value and acceptability increased probing more into trust-worthy and responsible AI to comply with stake-holder's expectations. This indispensable upward trend of more and more AI-integrated technologies is in harmony with safety and security oflife forms and environment. Here, some typical case studies shedding light on benefits of xAI methods in Medical diagnosis and health care are incorporated.

Keywords:AI; evolution (1950-to-2023-Future); eXplainable AI (xAI); interpretable/ Responsible/ Trustworthy AI; DARPAand NSF targets; Industry (Health, Defence, research requirements)-Deep architectures; CNN; Capsule Nets; ALEX; TRANSFORMERS; Machine Learning; Deep Learning—Supervised/unsupervised/Self-supervised data; Figures; Images, Tables, Scripts, Numerical values;

Artificial Intelligence [1950-2023 ...]

xAI.Med.		2022-62

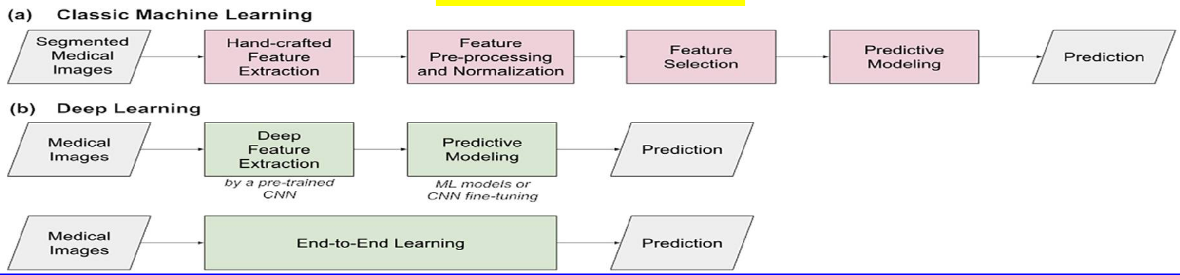
Timeline of AI, machine learning and deep learning



Narrow AI	A concept used to describe AI systemsthat are capable of handling a single or limited task
General AI	Hypothetical wisdom of AI systems capable of comprehending or learning any intelligent activity a human can perform
Super AI	AI that exceeds human intelligence and skills
Omni AI	Universal wisdom (pra-gyanam) comprising everything and anything and may be like Nature's intelligence.

xAI.Med.		2021-110

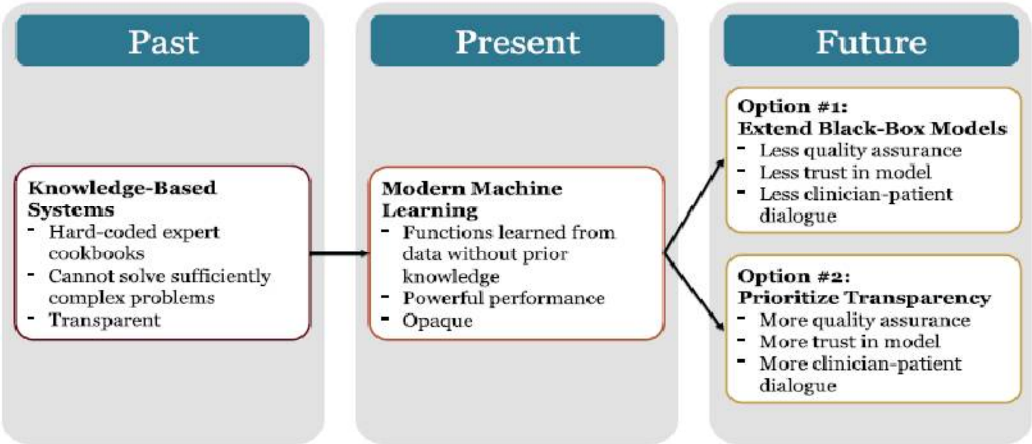
Architecture and workflow



- a) Classic machine learning, with the various processing steps involving hand-crafted features such as in radiomics
- B) Deep learning considering either deep medical image feature extraction or end-to-end learning

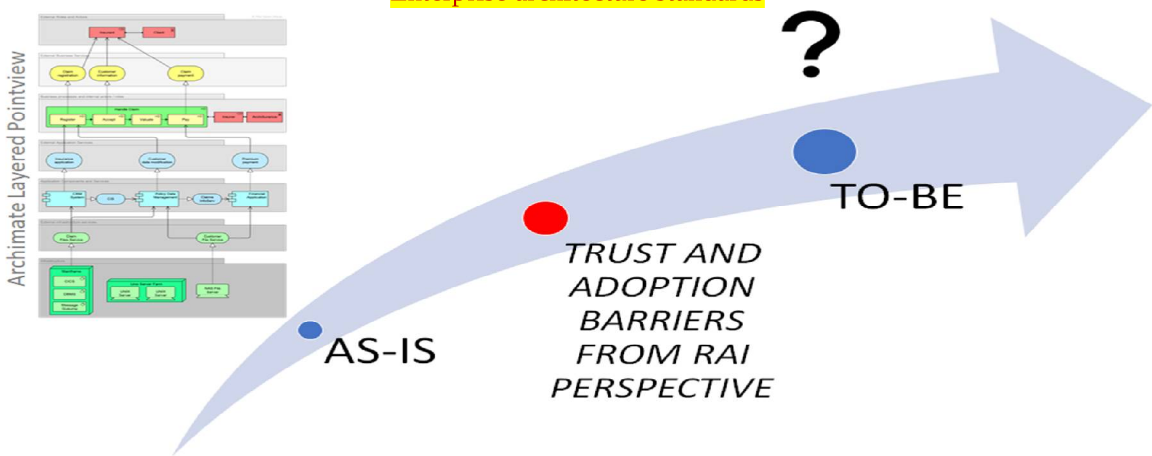
xAI.Med.		2022-70

AI in healthcare



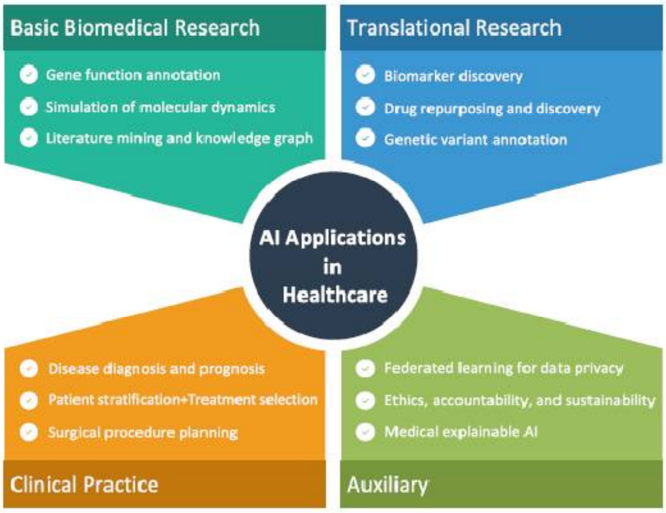
xAI.Med.	Health services	2022-49

Enterprise architecture standards



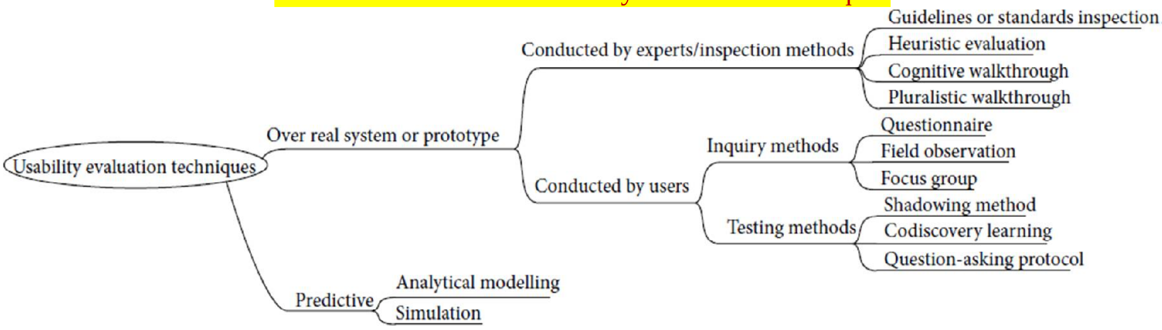
xAI.Med. | 2022-62

Non-exhaustive map of AI in healthcare applications



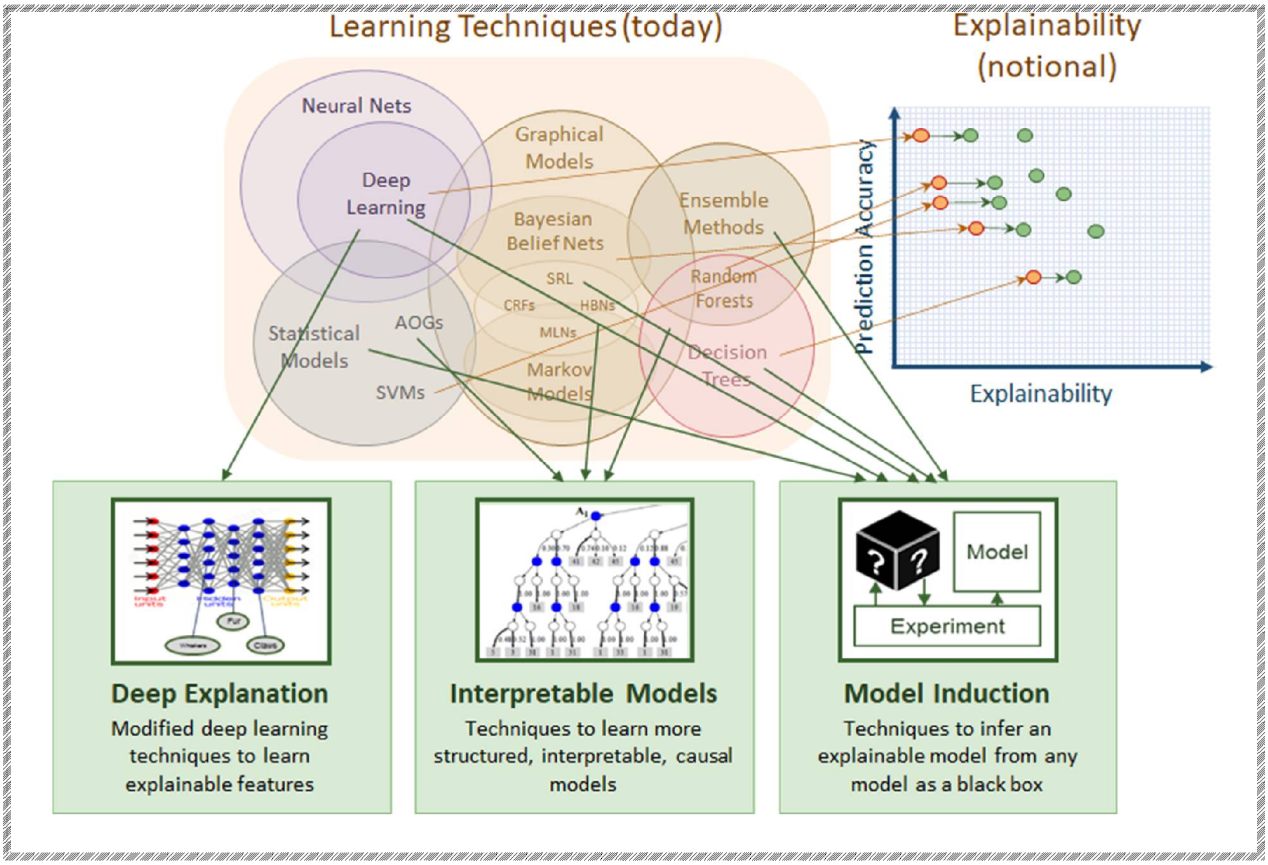
xAI.Med. | 2022-49

Classification of some usability evaluation techniques

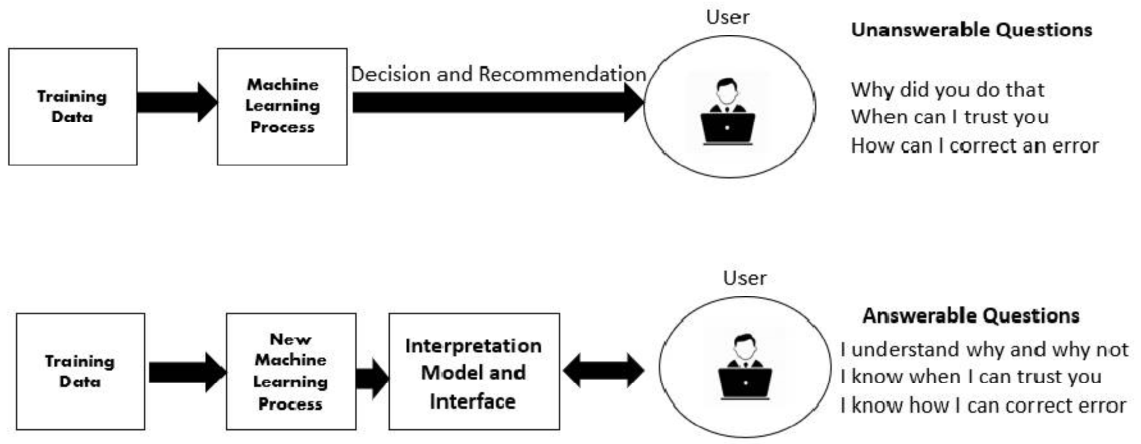


xAI.Med. | 2022-49

Explainable Models (DARPA)

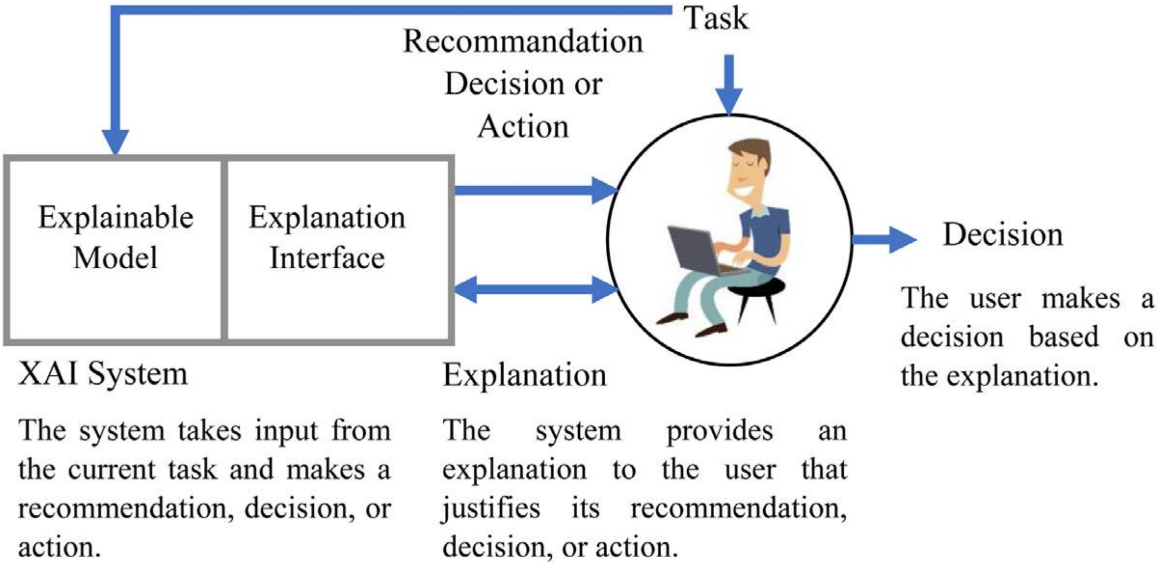


Explanation methods provide value to the end user



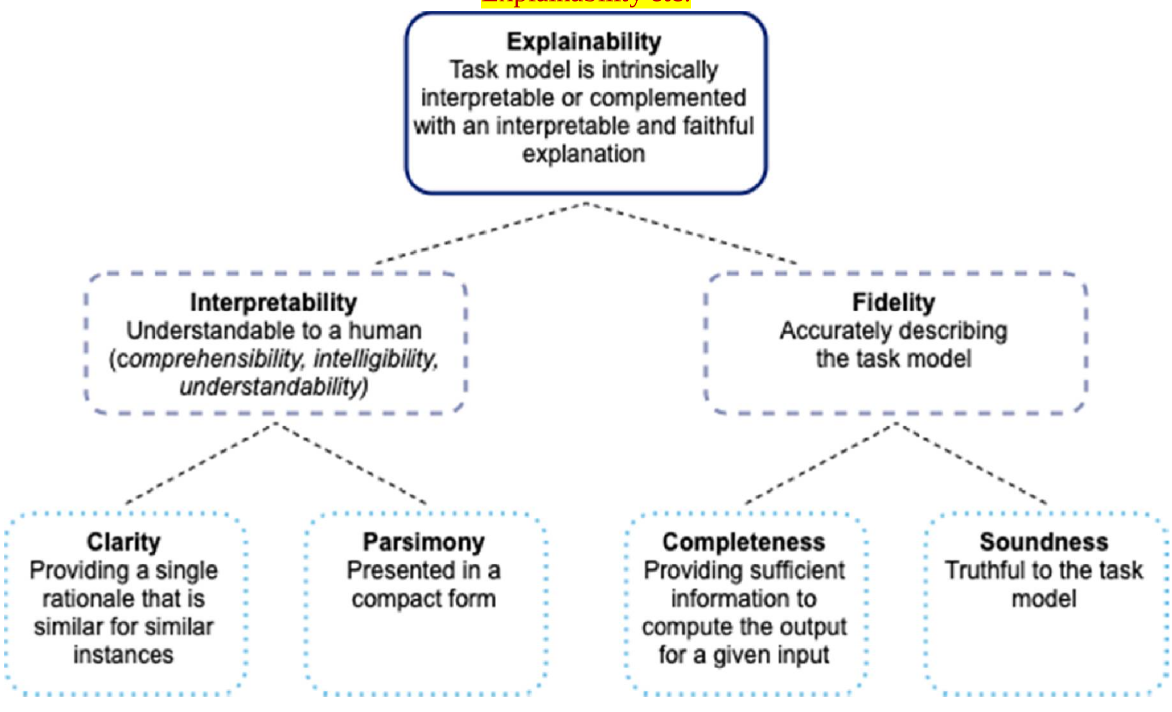
xAI.Med.		2021-20

xAI—Model and Expl



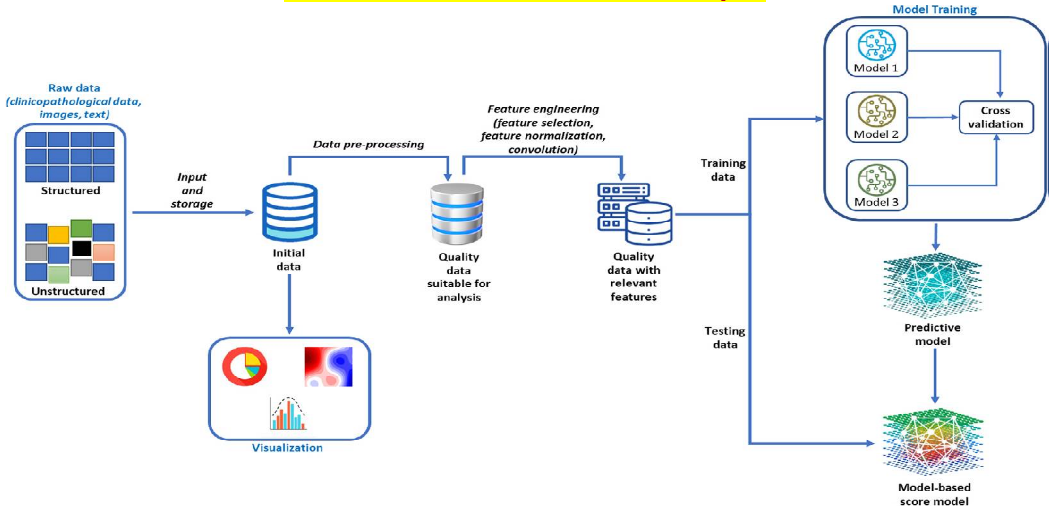
xAI.Med.		2021-19

Explainability etc.



xAI.Med.		2023-104

ML for Detection and Prediction Analysis



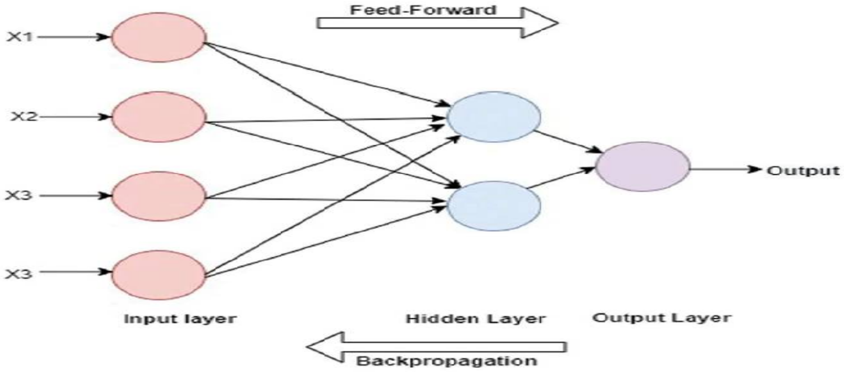
xAI.Med.		2022-49

Seven fundamental principles of privacy by design applied to the healthcare sector

Architectures | Black-box

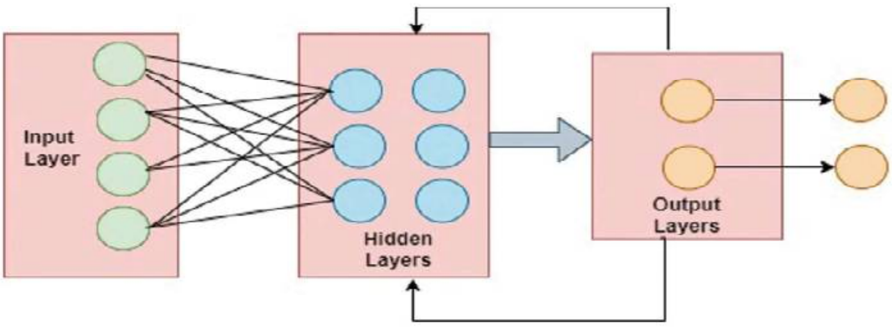
xAI.Med. | 2022-205

FF-SLP



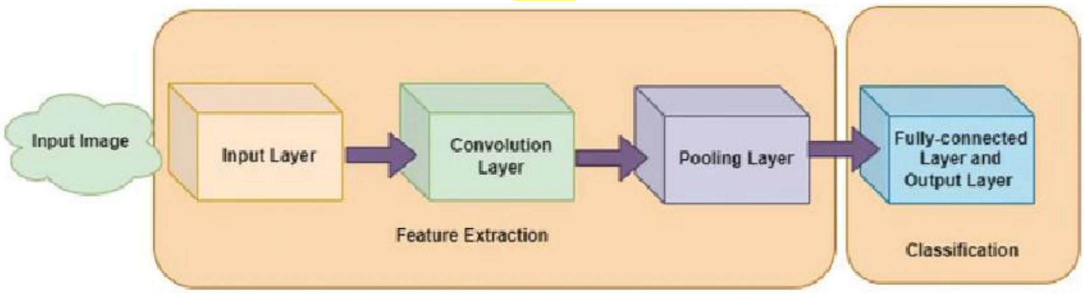
xAI.Med. | 2022-205

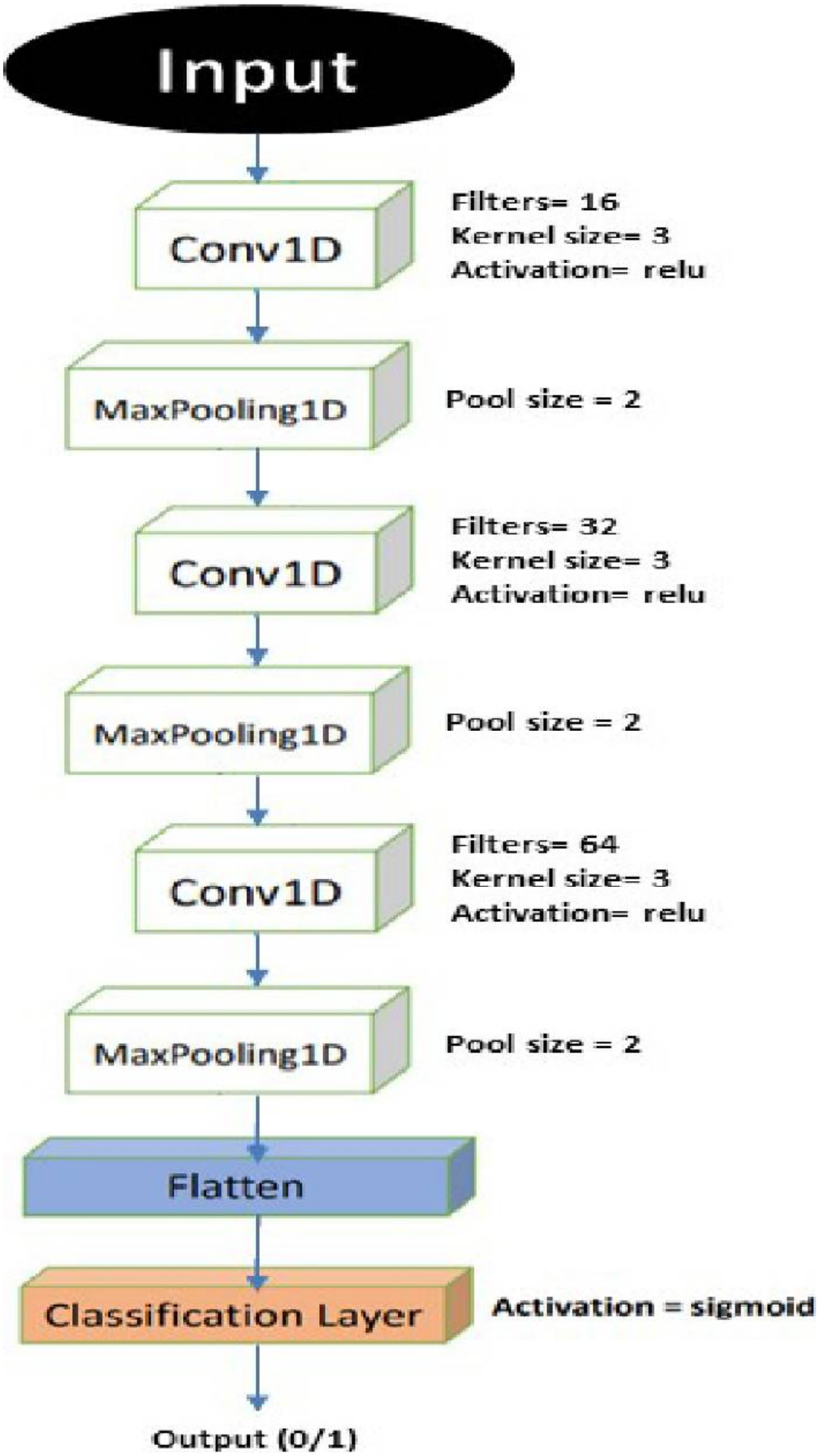
RecNN



xAI.Med. | 2022-205

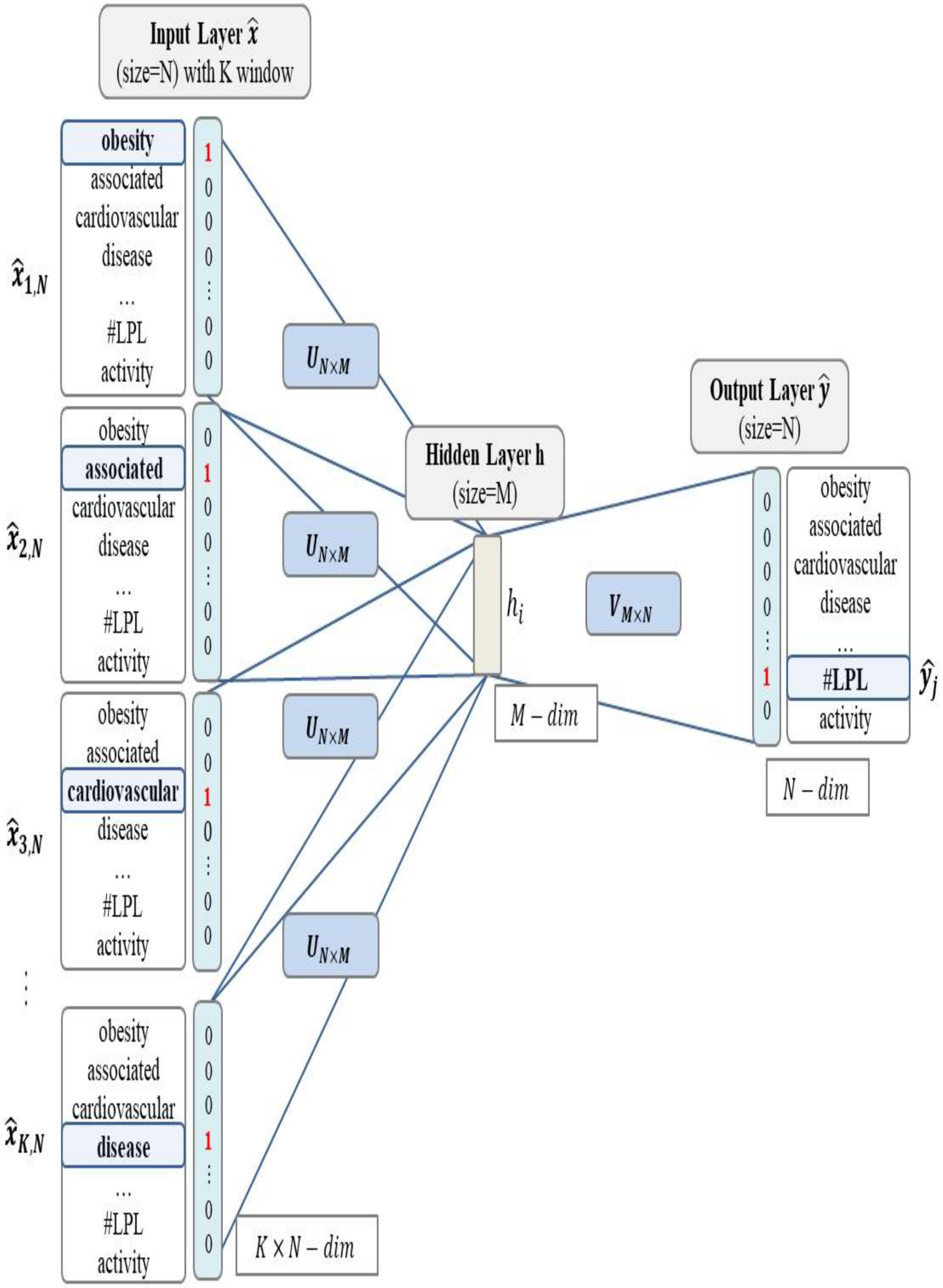
CNN





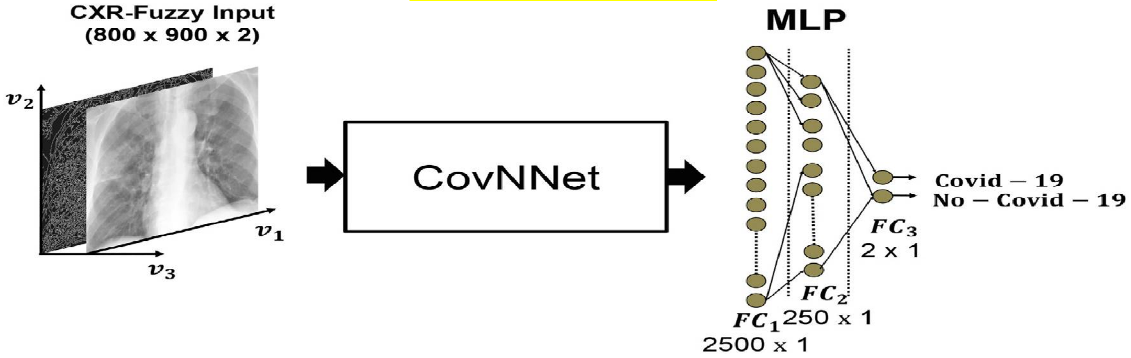
Input
Conv1D
Filters= 16
Kernel size= 3
Activation= relu
MaxPooling1D
Pool size = 2
Conv1D
Filters= 32
Kernel size= 3
Activation= relu
MaxPooling1D
Pool size = 2
Conv1D
Filters= 64
Kernel size= 3
Activation= relu
MaxPooling1D
Pool size = 2
Flatten
Classification Layer
Activation = sigmoid
Output (0/1)

Literature model structure (modified CBOW)



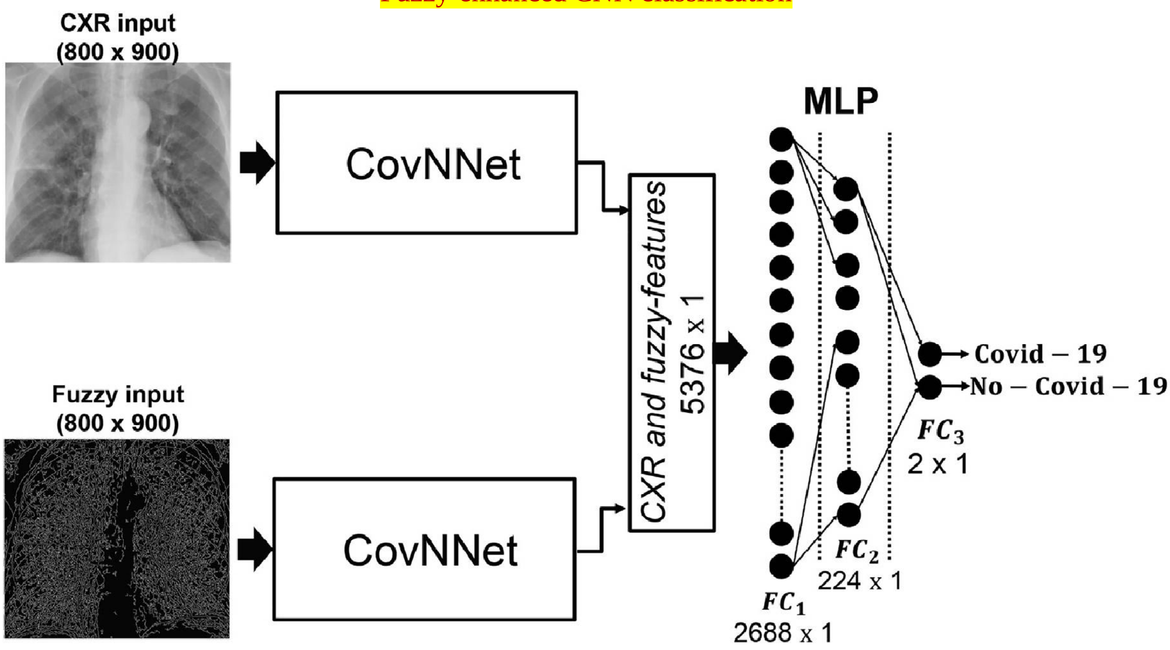
xAI.Med.	Fuzzy-CNN	2022-46

Covid-19 vs. No-Covid-19.



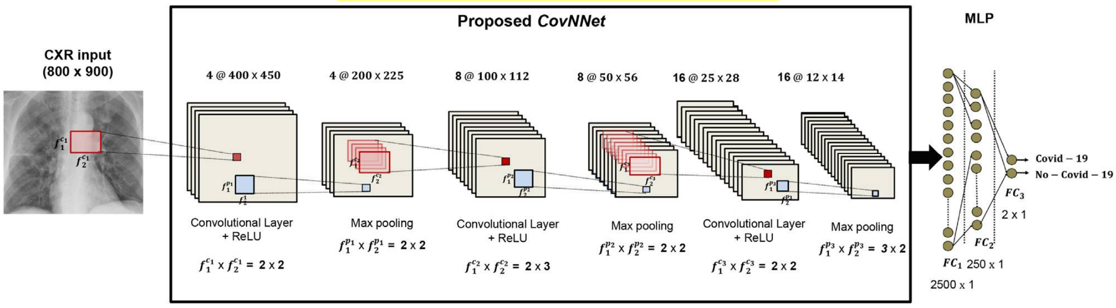
xAI.Med.	Covid-19 vs. No-Covid-19	2022-46

Fuzzy-enhanced CNN classification

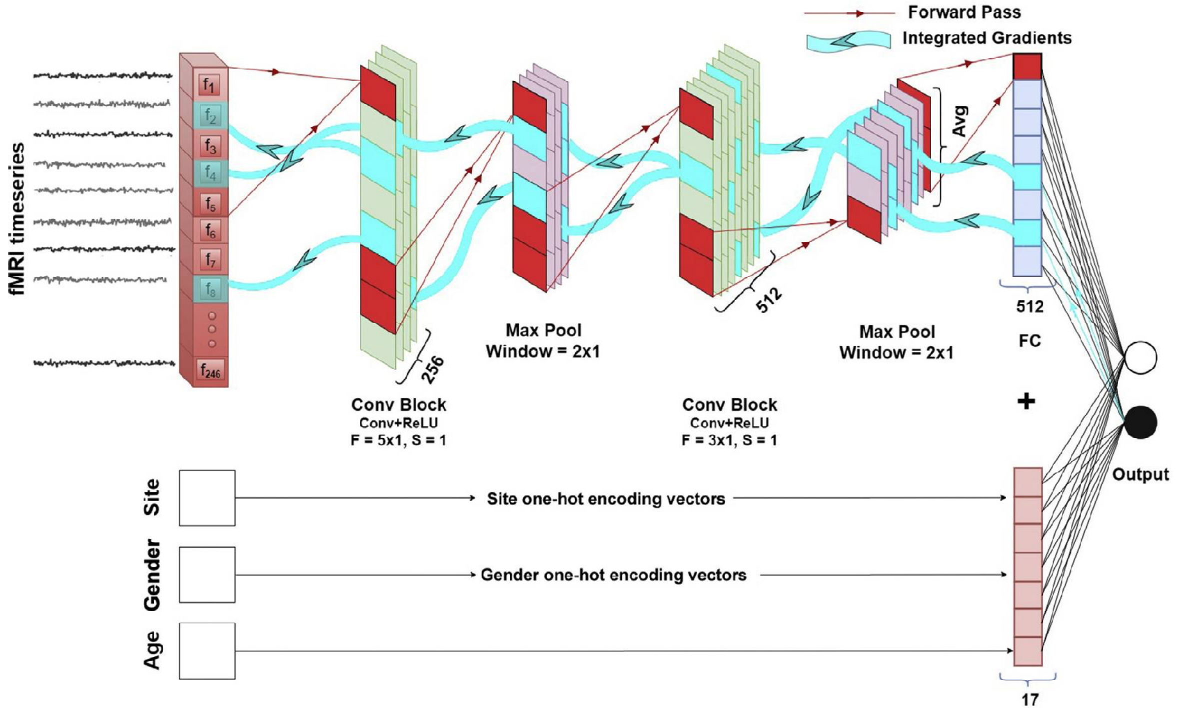


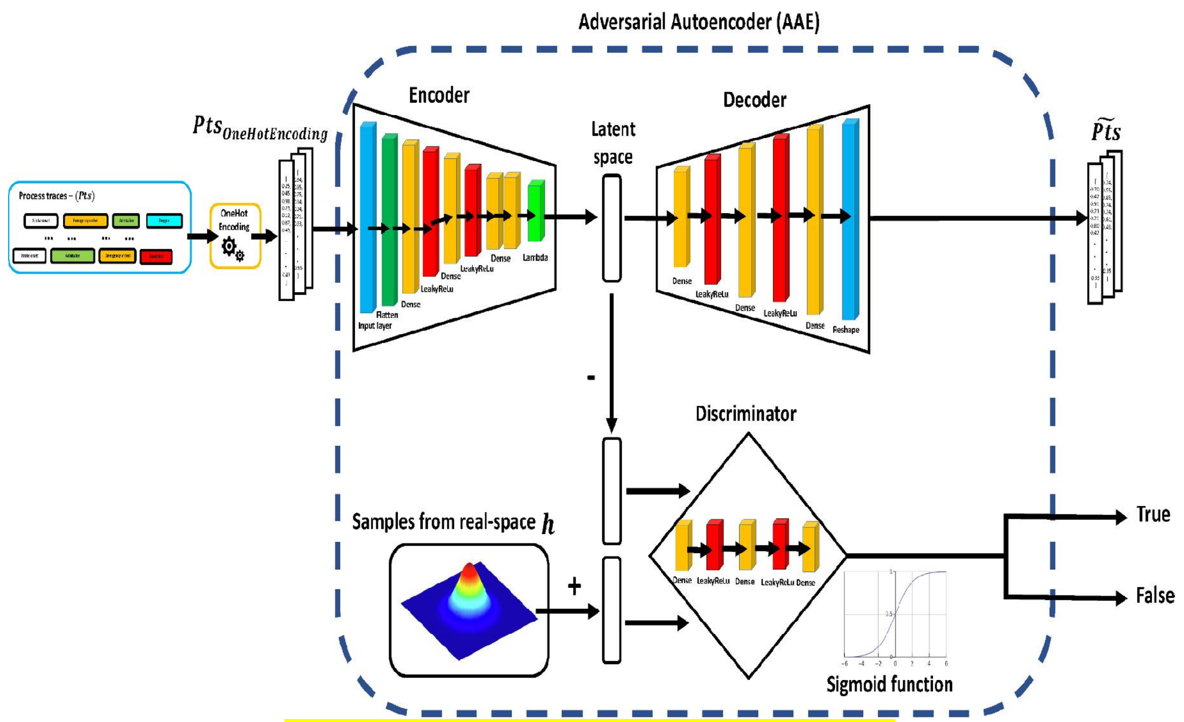
xAI.Med.	**Covid-19 vs. No-Covid-19 pneumonia classification**	2022-46

Covid-19 vs. No-Covid-19 pneumonia



Spatiotemporal CNN

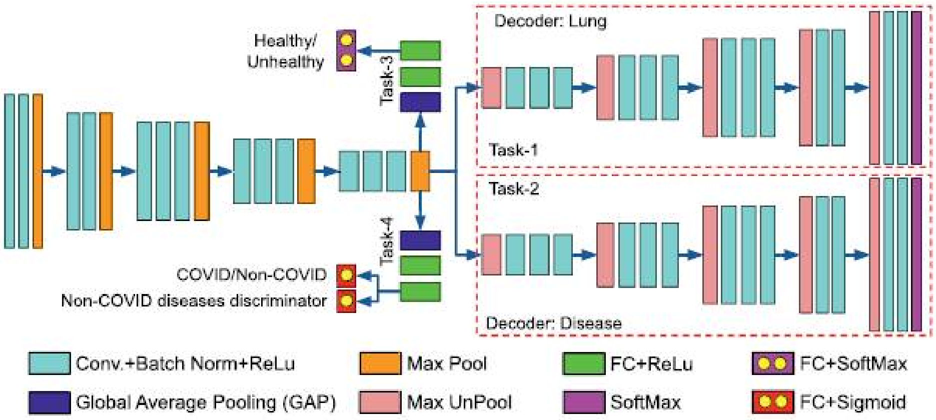




AAE architecture for trace neighbourhood generation

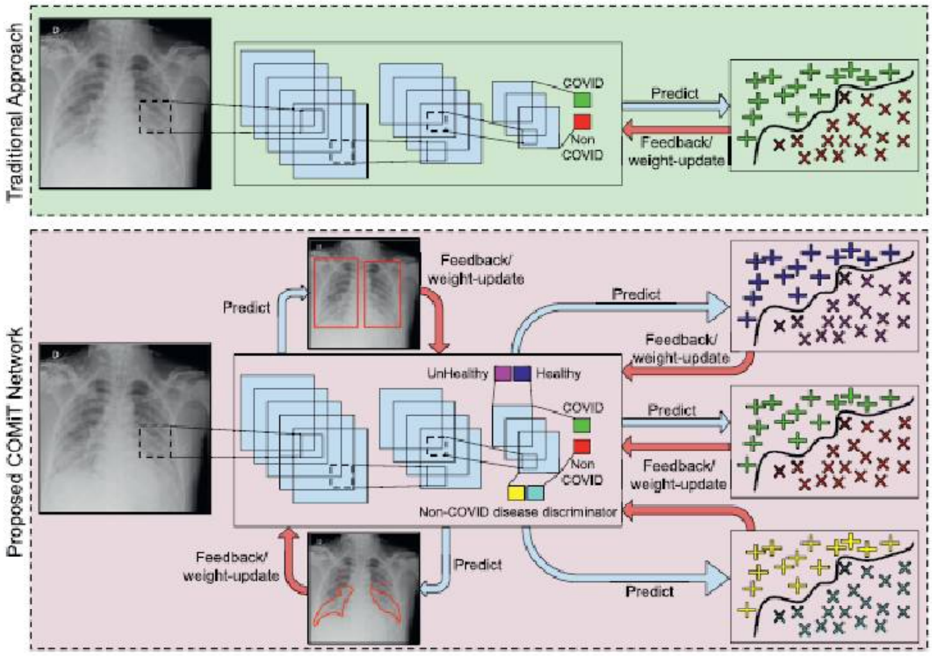
xAI.Med.	COVID-19	2022-75

Encoder-Decoder architecture



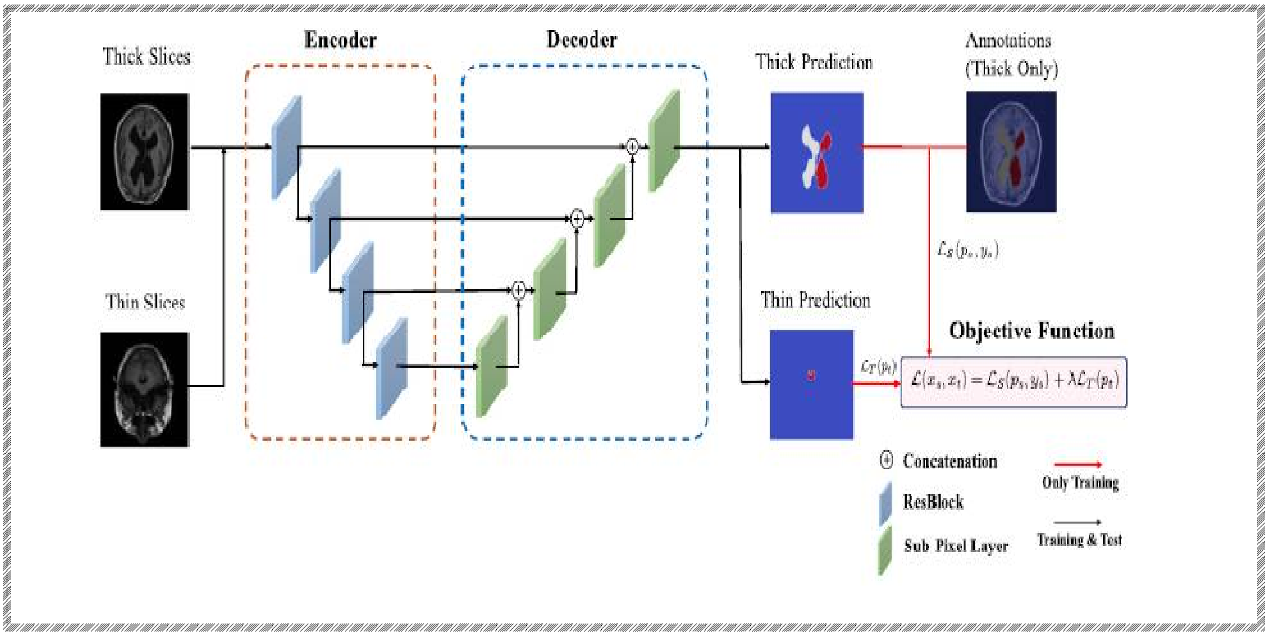
xAI.Med.	COVID-19	2022-75

COMiT



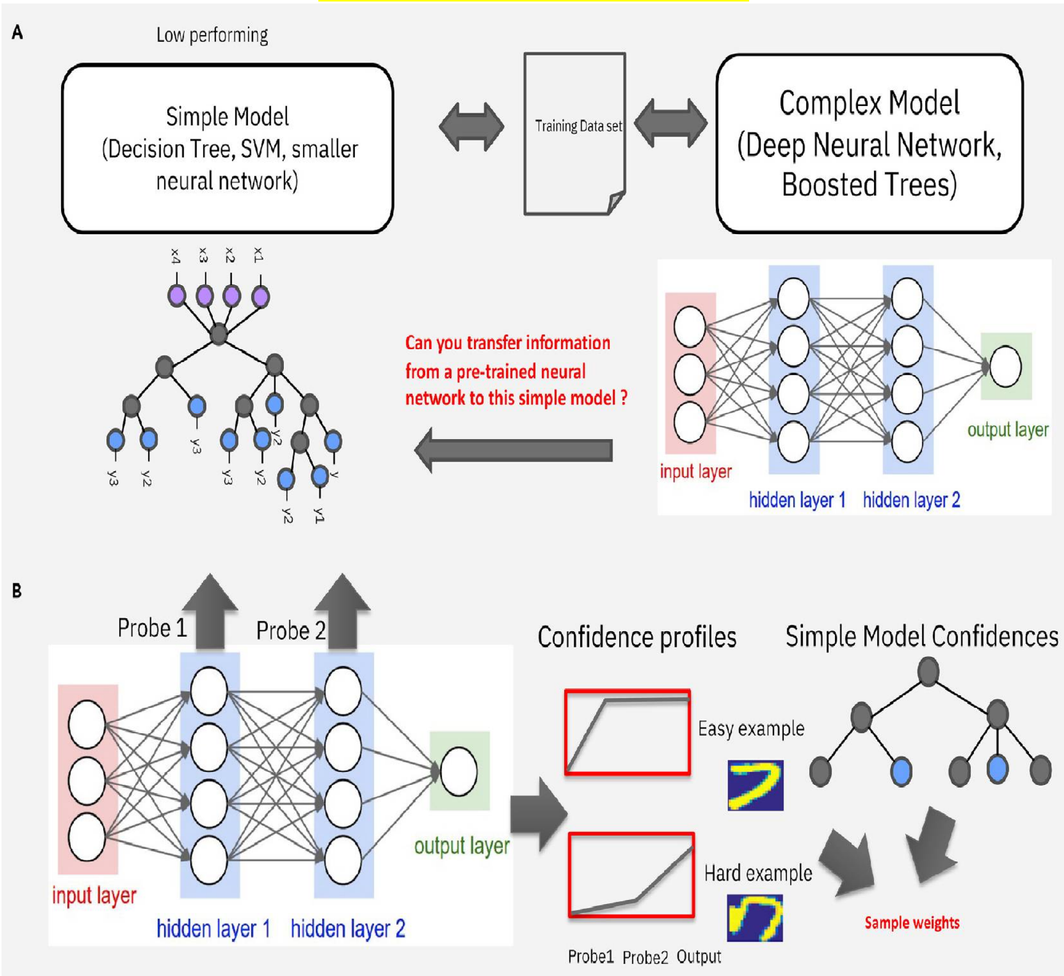
xAI.Med.	Segmentation	2022-62

XAI model for explainable segmentation

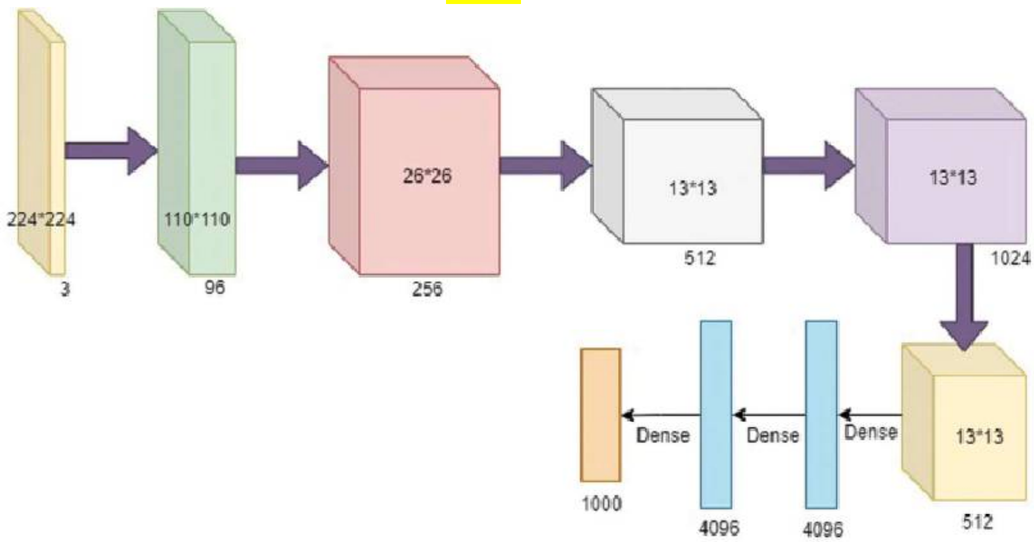


xAI.Med.	Number 7 from the MNIST	2022-36

Knowledge transfer for explainability

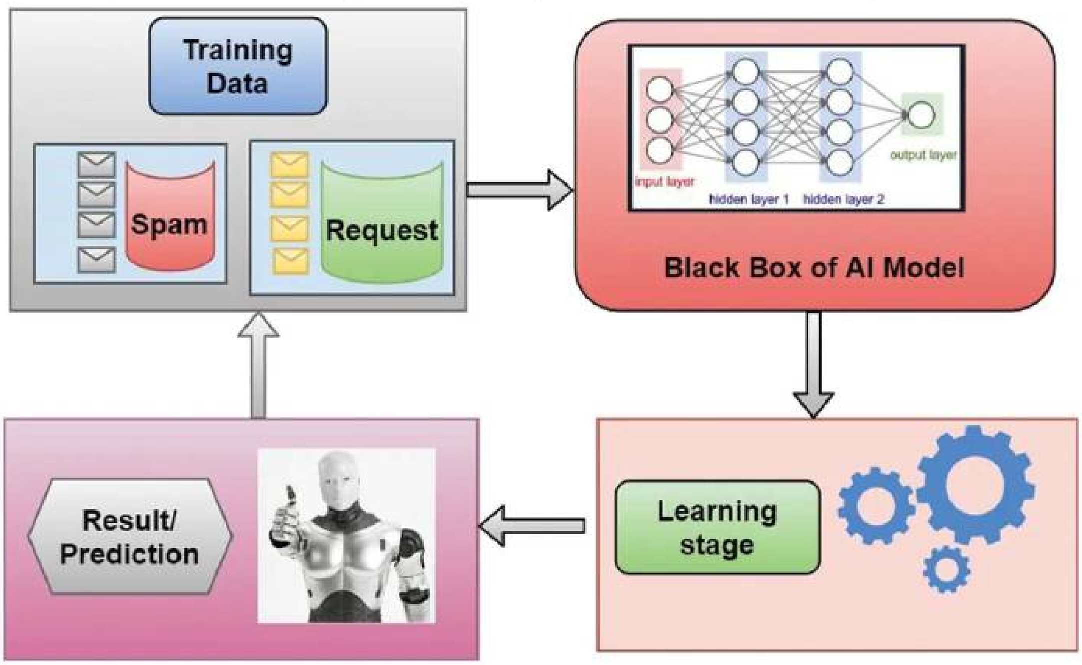


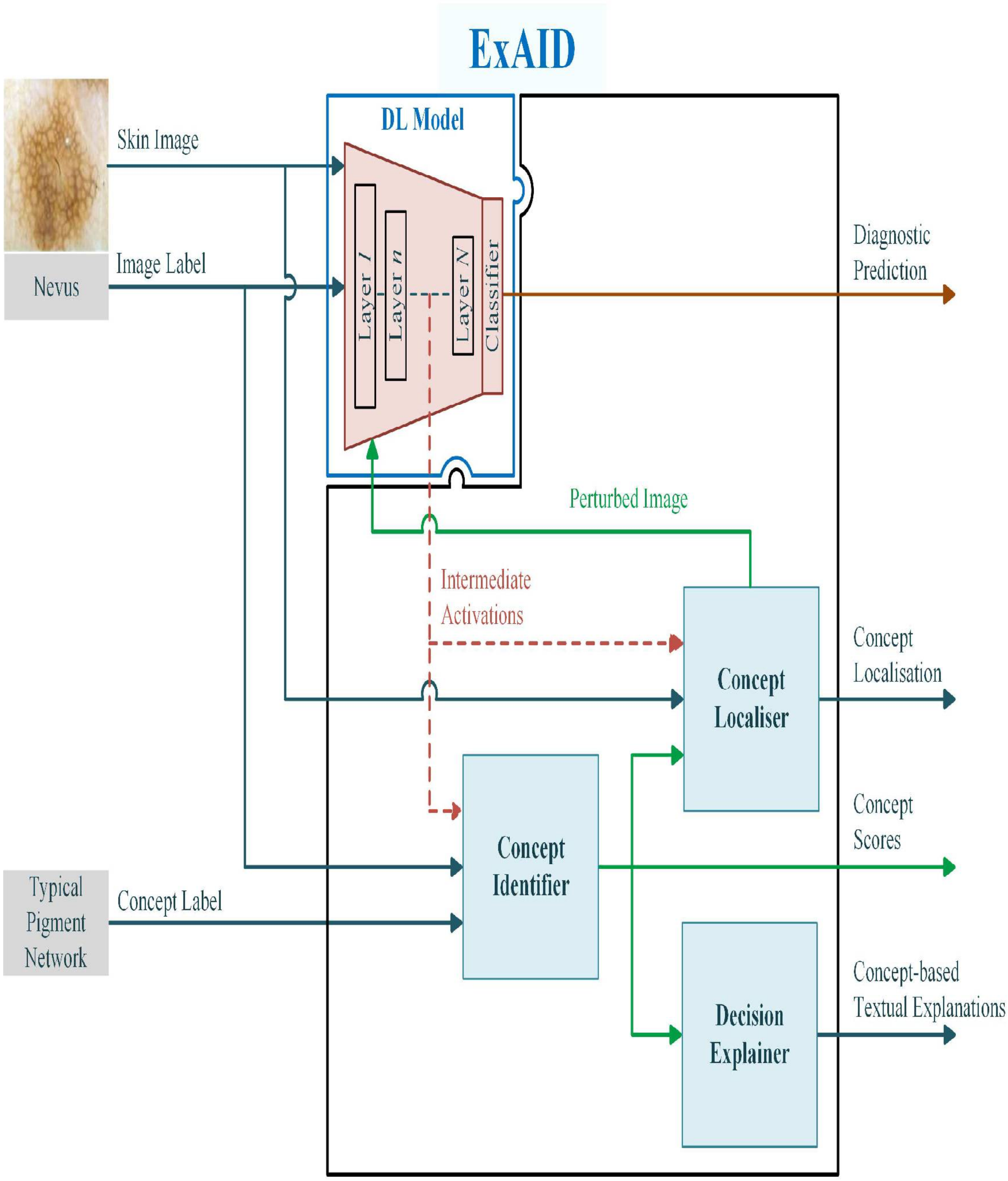
ZFNet



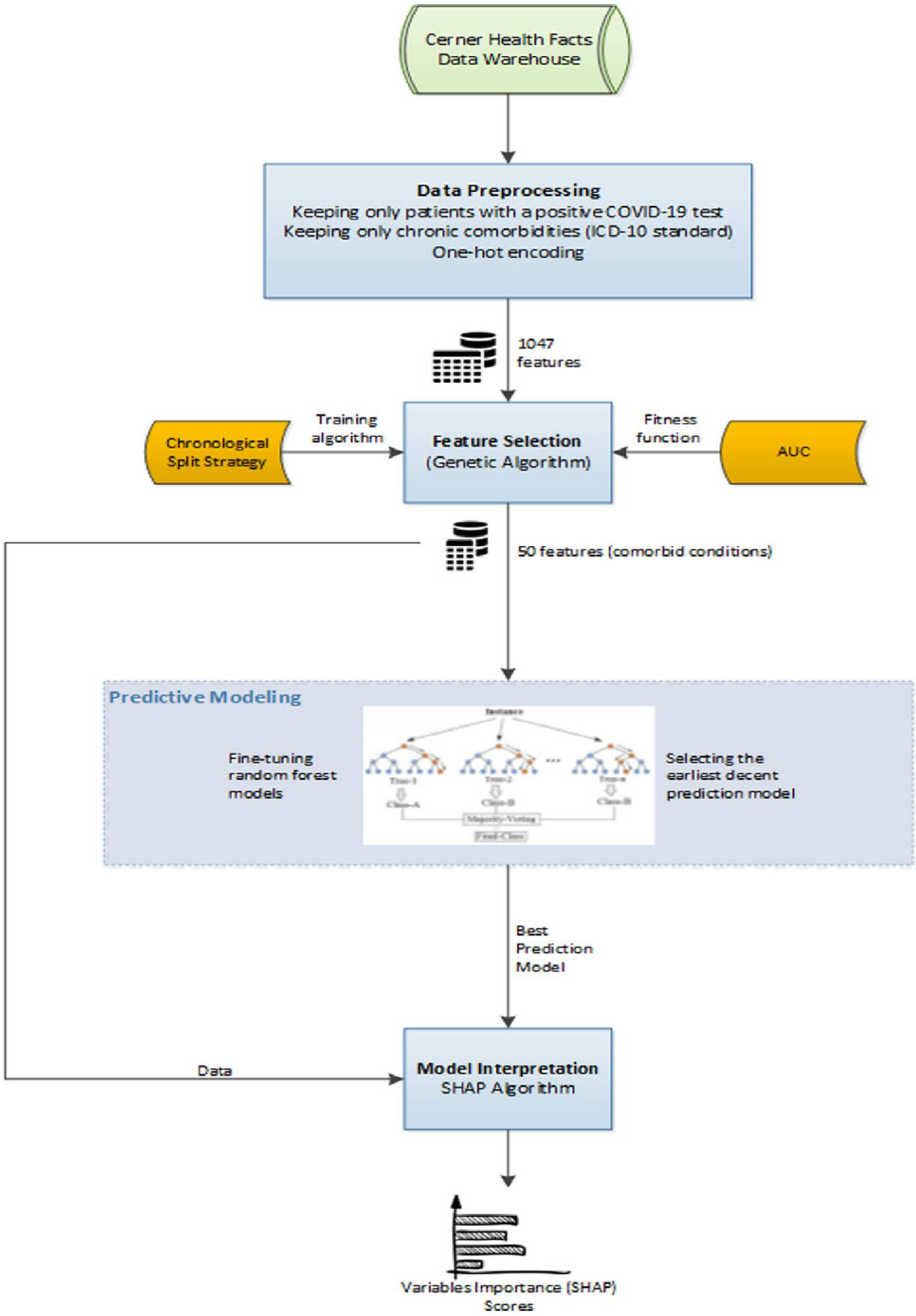
xAI.Med. 2022-205

AI-Classification




ExAID
DL Model
Skin Image
Image Label
Nevus
Layer 1
Layer n
Layer N
Classifier
Diagnostic
Prediction
Perturbed Image
Intermediate
Activations
Concept
Localiser
Concept
Localisation
Concept
Identifier
Concept
Scores
Typical
Pigment
Network
Concept Label
Decision
Explainer
Concept-based
Textual Explanations

xAI-Predictive modelling

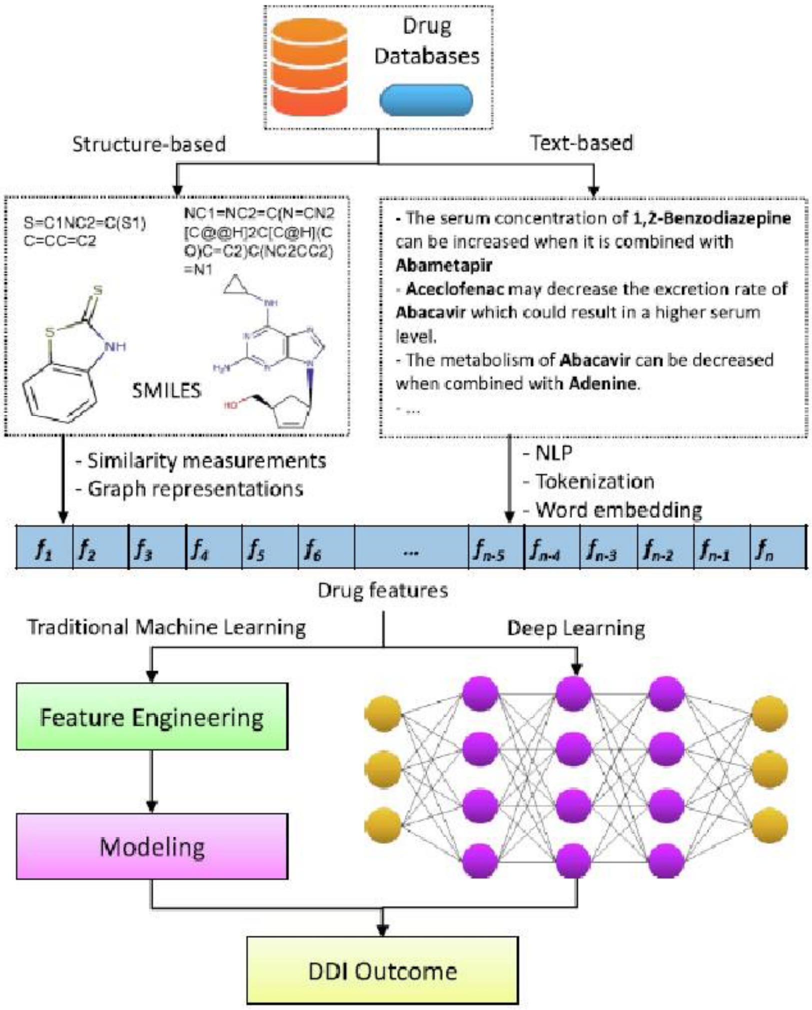


xAI.Med.		2022-32

xCNN-ECG

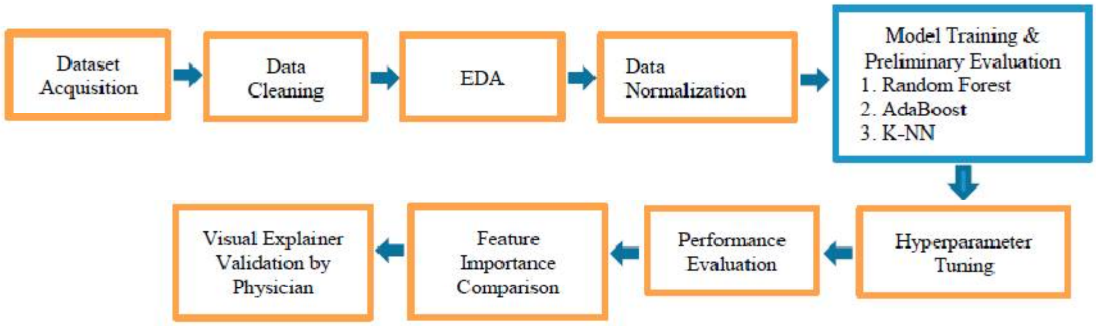
xAI.Med.	Drug-Drug-Interaction-- prediction	2022-39

Traditional ML and DL for DDIs prediction

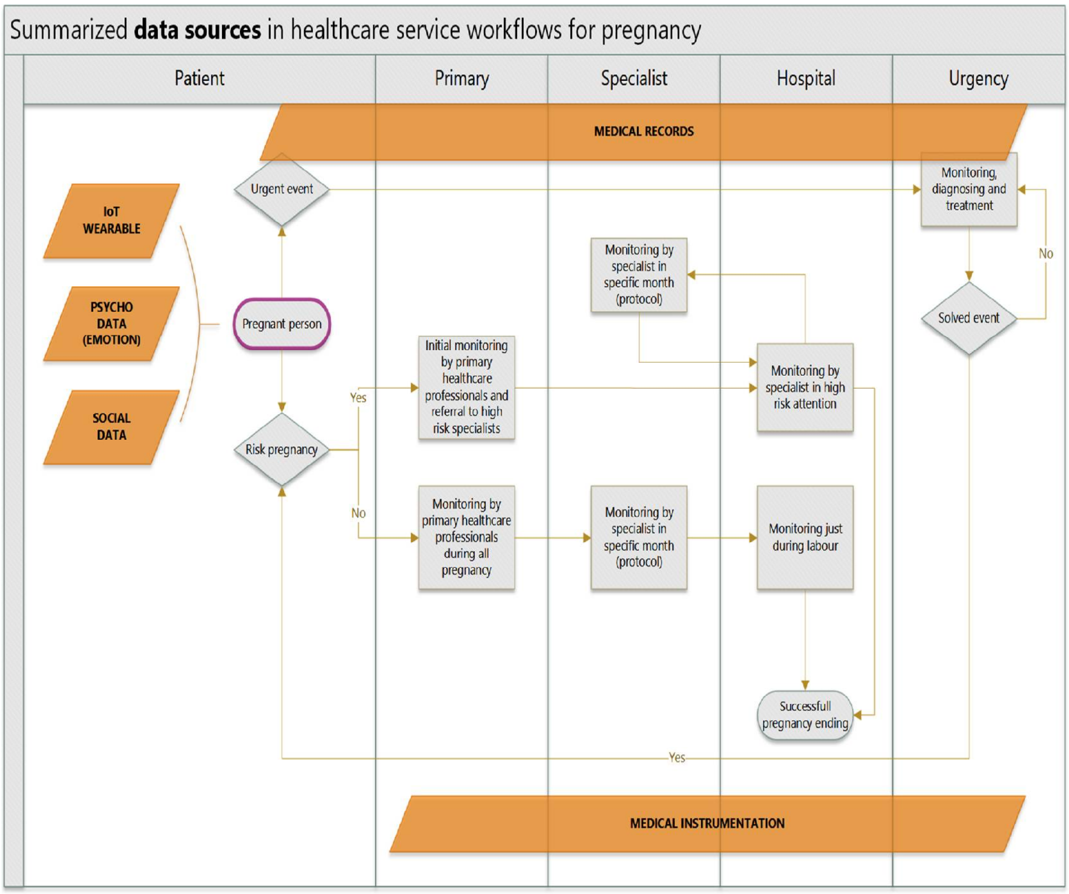


xAI.Med.		2022-40

Hospital Readmission Prediction Methodology



xAI.Med.	pregnancy	2022-49

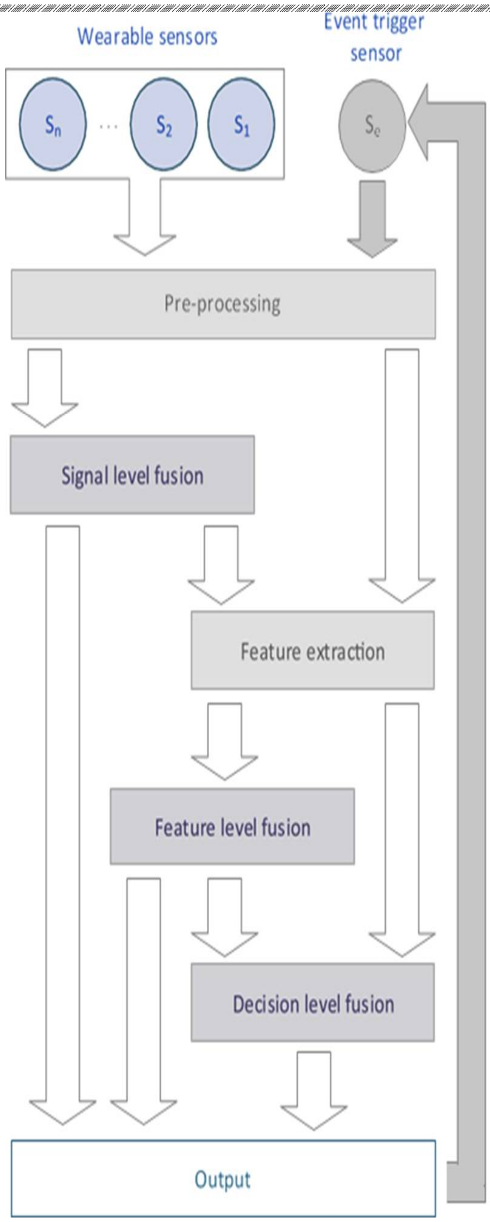


Fusion Data

xAI.Med.	Wearable health monitoring systems	2022-49

Data fusion architecture

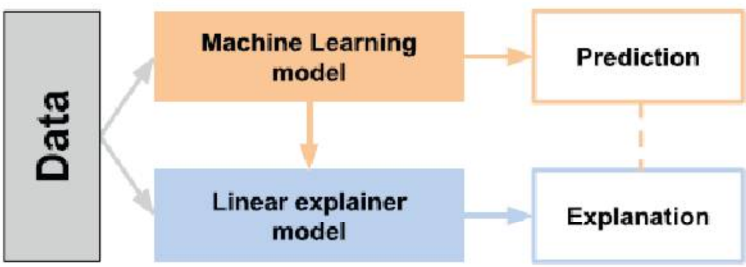
Summary of data fusion algorithms



Level	Algorithm
Signal/ sensor	Weighted averages
	The Kalman filter (KF)
	Particle filtering (PF)
Feature (non-parametric)	k-Nearest Neighbour (k-NN)
	Decision Trees (DT)
	Support Vector Machines (SVM)
	Artificial Neural Network (ANN) and Deep Learning
Feature (parametric)	Gaussian mixture model (GMM)
	k-Means
Decision	Bayesian inference
	Fuzzy logic
	Topic models

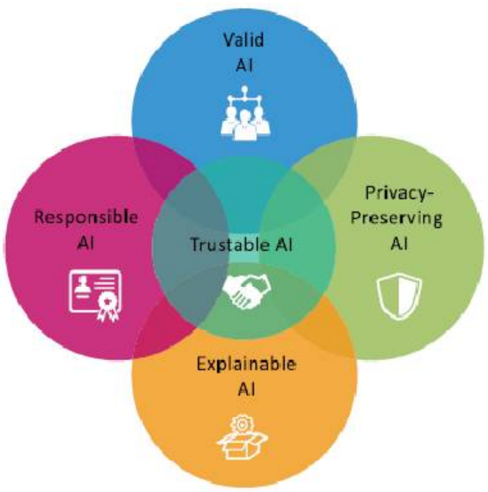
xAI.Med.	2022-45

Interpretability used to explain machine learning models



$AI	
XAI	RAI
Explainable AI	Responsible AI
PPAI	TAI
Privacy-Preserving AI	Trustable / TrustworthyAI

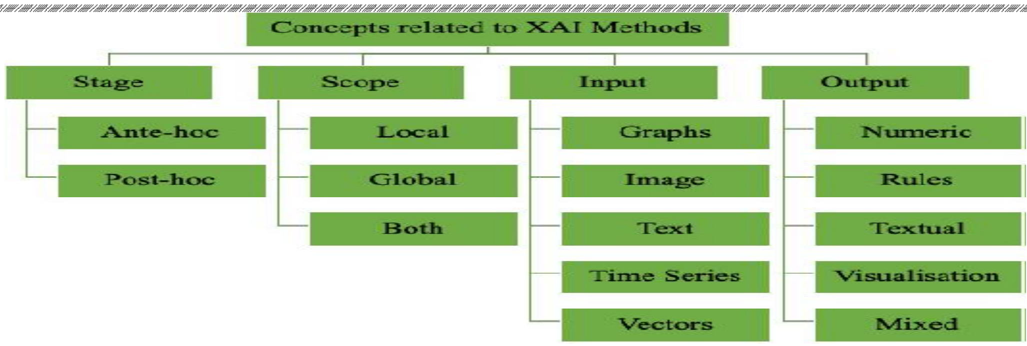
xAI.Med.		2022-62



Trustable AI or Trustworthy AI	EU General Data Protection Regulation (GDPR)

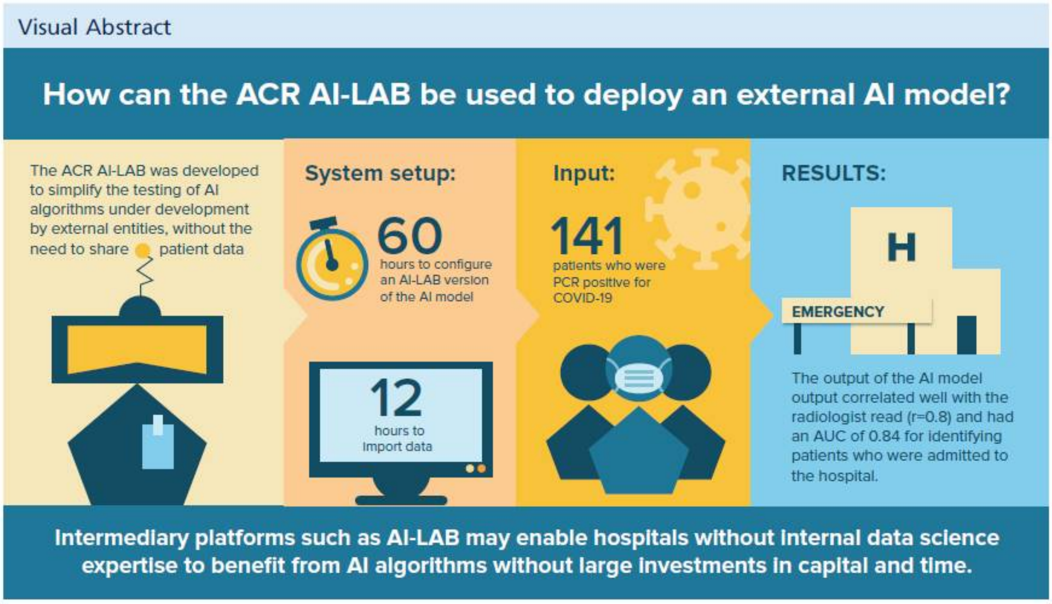
Probes xAI

AI.Med.		2022-205
Different concepts in developing methodologies for XAI		



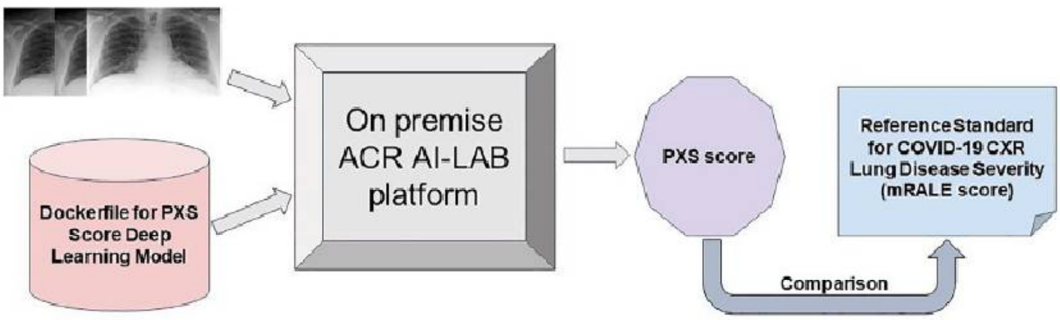
xAI.Med.		2022-202

ACR-AI Lab

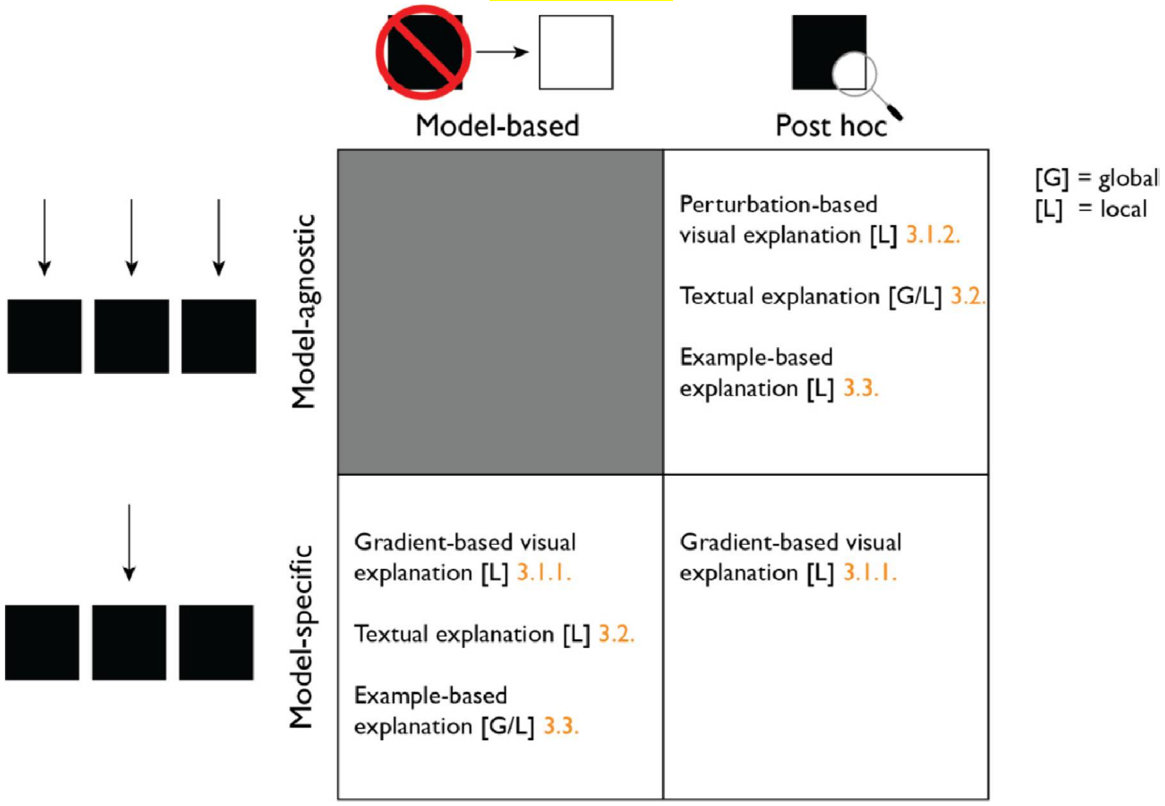


xAI.Med.		2022-202

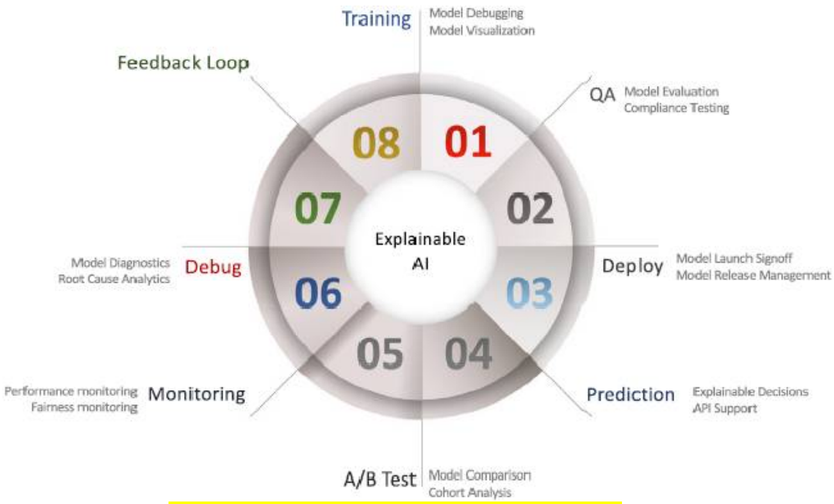
Chest X-ray radiograph--- Lung Edema -- Pulmonary X-ray Severity



XAI framework



Training, quality assurance (QA), deployment, prediction, split testing (A/B test) monitoring,and debugging



Feedback loop of the XAI development

Pros (+) and cons (-) of XAI

Pros and cons of XAI techniques. Pros are depicted by +, cons by -. The letters in the column Open source (original paper) refer to the URL below the table.

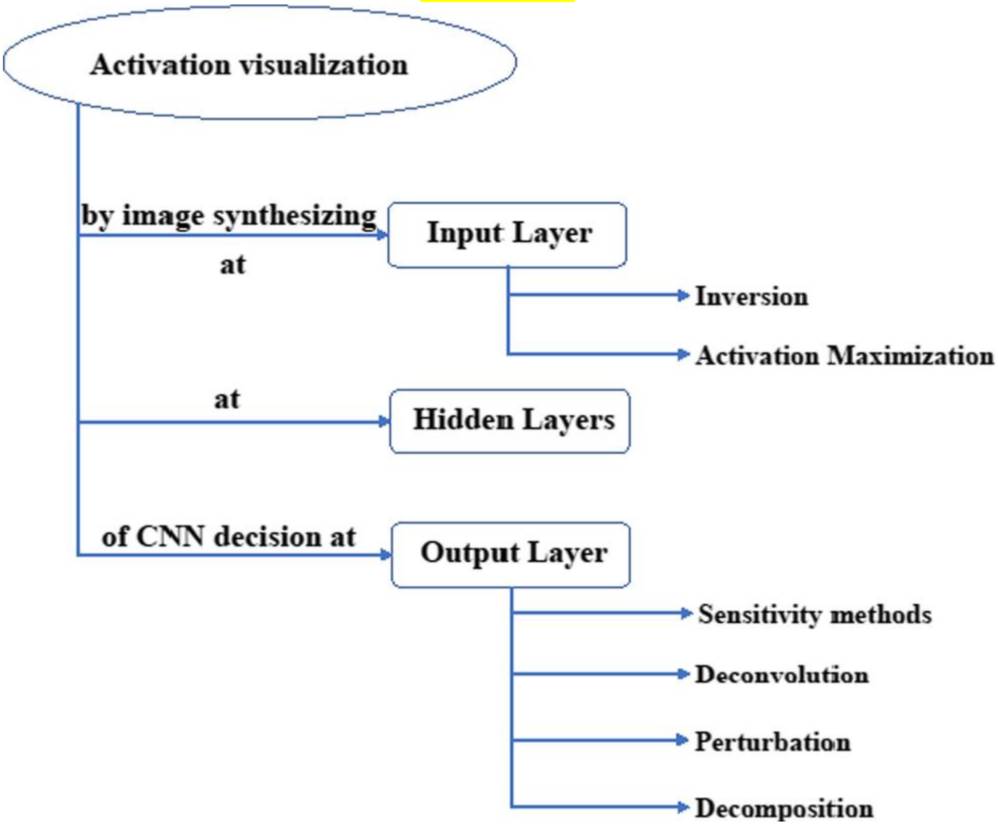
Technique	Ease of use	Validity	Robustness	Computational needs	No fine-tuning required	Open-source (original paper)	Open-source (captum.ai)
Visual explanation							
Backpropagation-based approaches							
Backpropagation	+	-	+	-	+	-	+
Deconvolution	+	n.t.	n.t.	-	+	-	+
Guided backpropagation	+	-	inc.	-	+	-	+
Class activation mapping (CAM)	+	n.t.	-	-	+	a	-
Gradient-weighted class activation mapping (Grad-CAM)	+	+/-	-	-	+/-	b	+
Layer wise relevance propagation (LRP)	|	n.t.	|		|/		|
Deep SHapley Additive exPlanations (Deep SHAP)	+	n.t.	n.t.	-	+/-	c	+
Trainable attention	+/-	n.t.	n.t.	+	-	d	-
Perturbation-based approaches							
Occlusion sensitivity	+	n.t.	-	+	-	-	+
Local Interpretable Model-agnostic Explanations (LIME)	+	n.t.	n.t.	+	-	e	+
Meaningful Perturbation	+	n.t.	n.t.	+	-	f	-
Prediction difference analysis	+	n.t.	n.t.	+	-	g	-
Textual explanation							
Image captioning	+/-	n.t.	n.t.	+	-	-	-
Image captioning with visual explanation	+/-	n.t.	n.t.	+	-	h	-
Testing with Concept Activation Vectors (TCAV)	+	n.t.	n.t.	n.t.	+/-	i	-
Example-based explanation							
Triplet networks	+/-	n.t.	n.t.	+	-	j	-
Influence functions	+	n.t.	n.t.	n.t.	+/-	k	-
Prototypes	+/-	n.t.	n.t.	+	-	l	-

n.t. = not tested by studies on that criterion.
inc. = inconclusive results between studies on that criterion.
a https://github.com/zhoubolei/CAM
b https://github.com/Cloud-CV/Grad-CAM
c https://github.com/slundberg/shap
d https://github.com/saumya-jetley/cd_ICLR18_LearnToPayAttention
e https://github.com/marcotcr/lime
f https://github.com/ruthcfong/perturb_explanations
g https://github.com/lmzintgraf/DeepVis-PredDiff
h https://github.com/zizhaozhang/tandemnet
i https://github.com/tensorflow/tcav
j https://github.com/eladhoffer/TripletNet
k https://github.com/kohpangwei/influence-release
l https://github.com/cfchen-duke/ProtoPNet

Global Information Tracker (git) hub

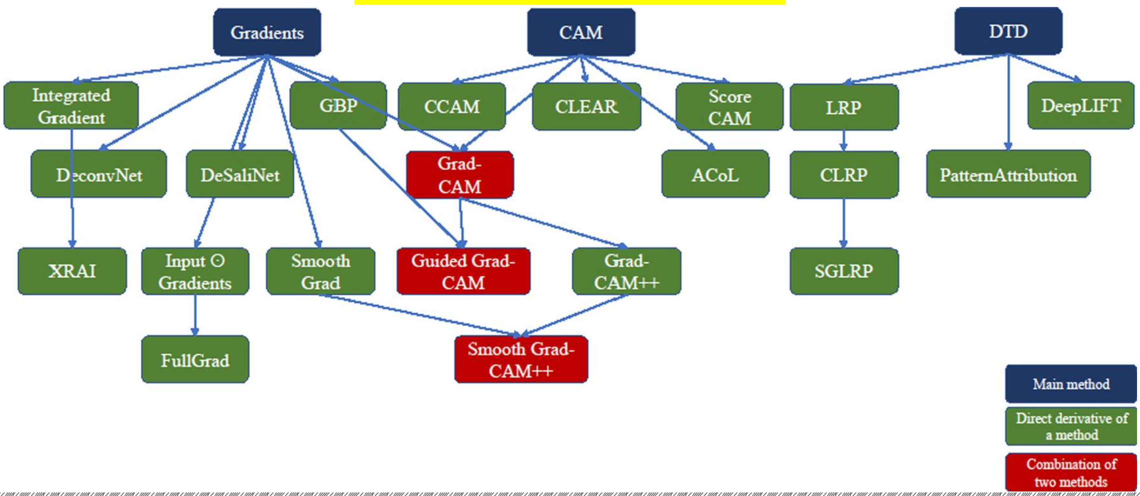
xAI.Med. | 2023-58

xAI-Methods



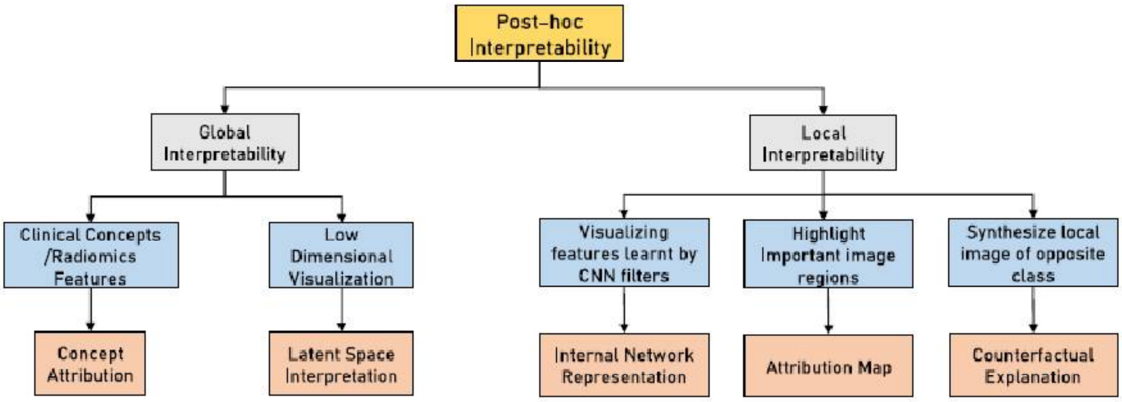
xAI.Med. | 2023-58

Post-hoc -xAI--visualisation Methods



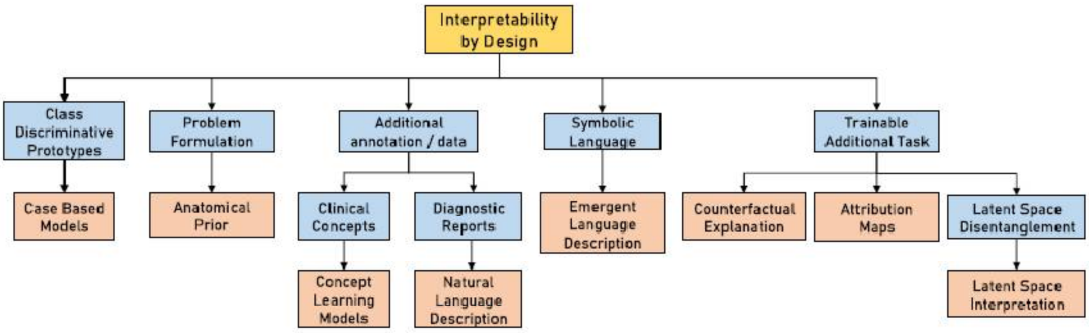
xAI.Med.		2022-68

Medical image analysis
DeepLrn—DeepArch--xMed



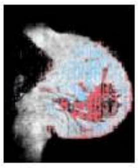
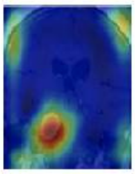
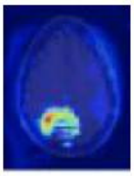
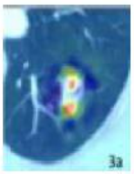
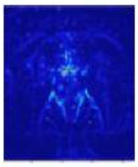
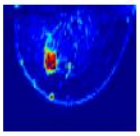
xAI.Med.		2022-68

Interpretability during the design process of the deep neural network

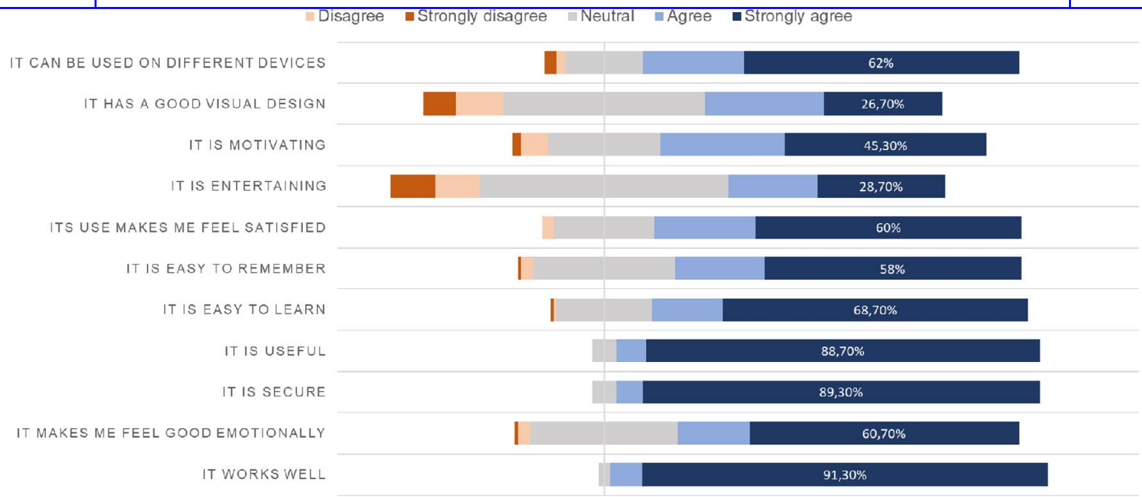


xAI.Med.		2022-208

Post-Hoc explanation techniques for cancer classification

Method	Results	Perturbation or Gradient	Local or Global	Model-Agnostic or Model-Specific
SHAP		Perturbation	Both	Model-Agnostic
CAM		Gradient	Local	Model-Specific
Grad-CAM		Gradient	Local	Model-Specific
PG-CAM		Gradient	Local	Model-Specific
Occlusion		Perturbation	Local	Model-Specific
Saliency Map		Gradient	Local	Model-Agnostic
Integrated Gradients		Gradient	Local	Model-Agnostic

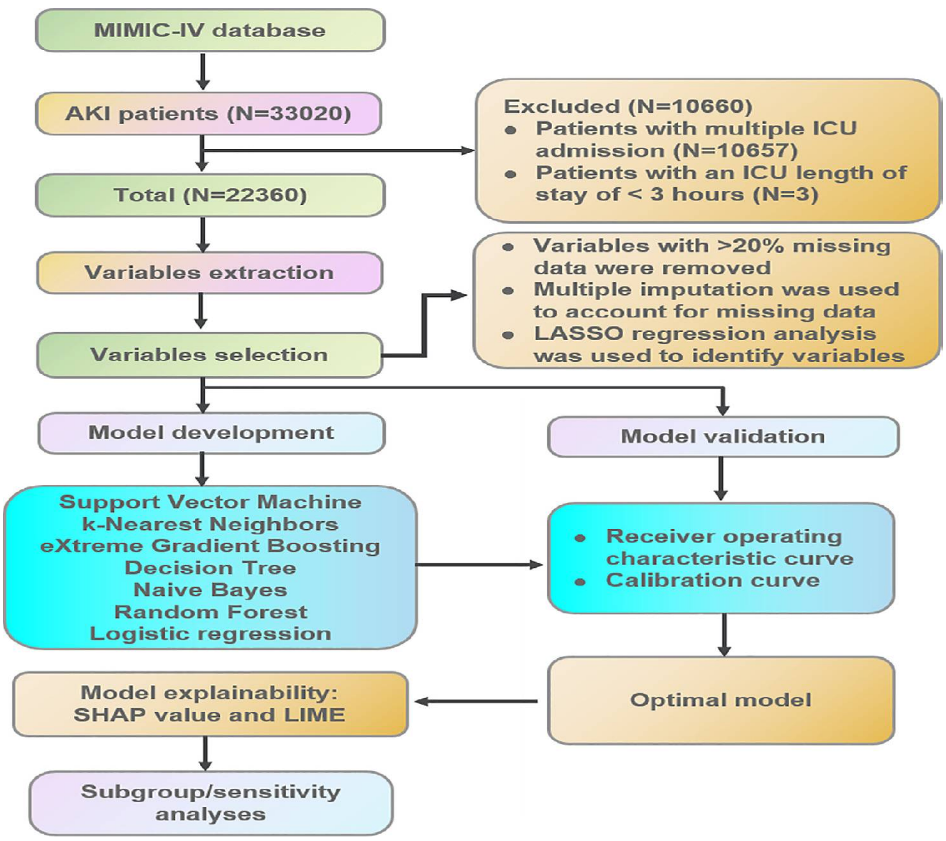
xAI.Med.	Valuableaspects applications	2022-49



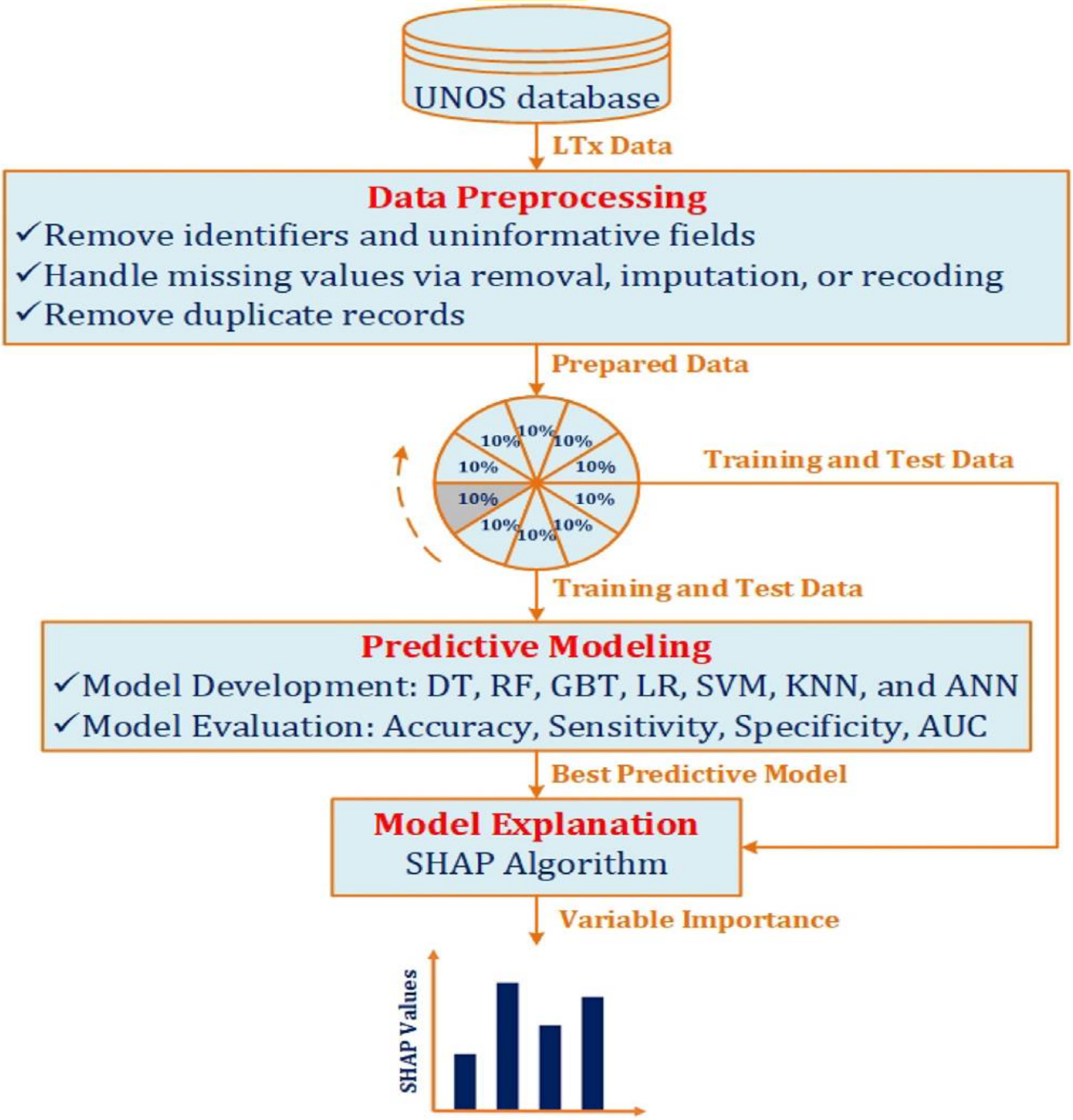
xAI.Med.	MIMIC(Medical Information Mart for Intensive Care)	2022-44

LASSO (Least absolute shrinkage and selection operator)

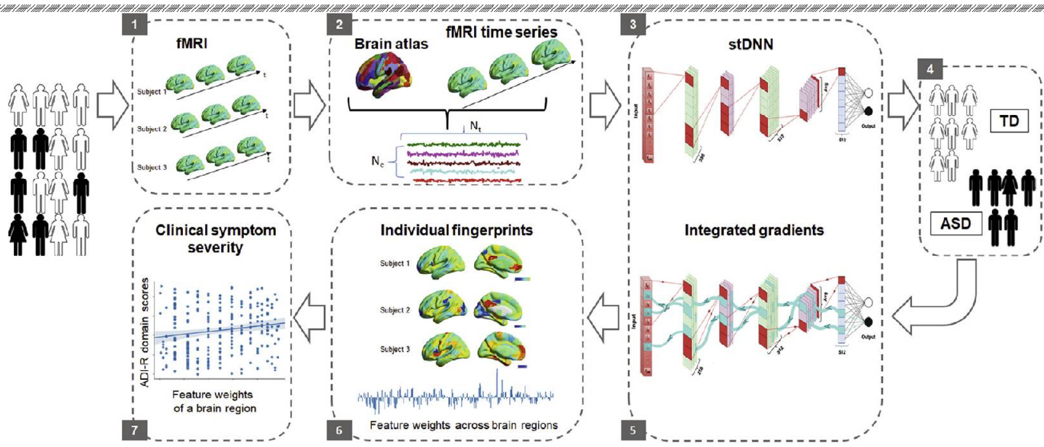
SHAP, SHapley Additive



SHAP Alg



To predict autism

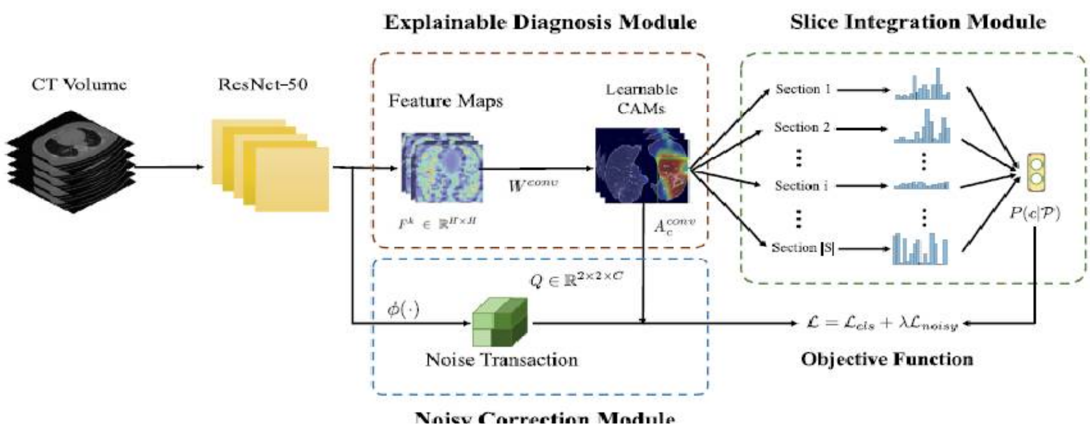


- Steps 1 and 2, data extraction;
- Steps 3 and 4, classification;
- Steps 5 and 6, feature identification, i.e., predictive feature weights (fingerprints) across brain regions;
- step 7, prediction of clinical symptom severity

.

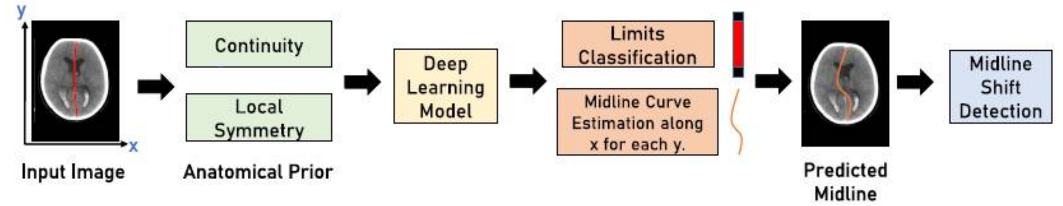
- ✓ ADI-R, Autism Diagnostic Interview-Revised;
- ✓ fMRI, functional magnetic resonance imaging;
- ✓ stDNN, spatiotemporal deep neural network;
- ✓ TD, typically developing

xAI.Med.	COVID-19	2022-62



P (*c* | p) :Probability of the patient who is COVID-19 infected

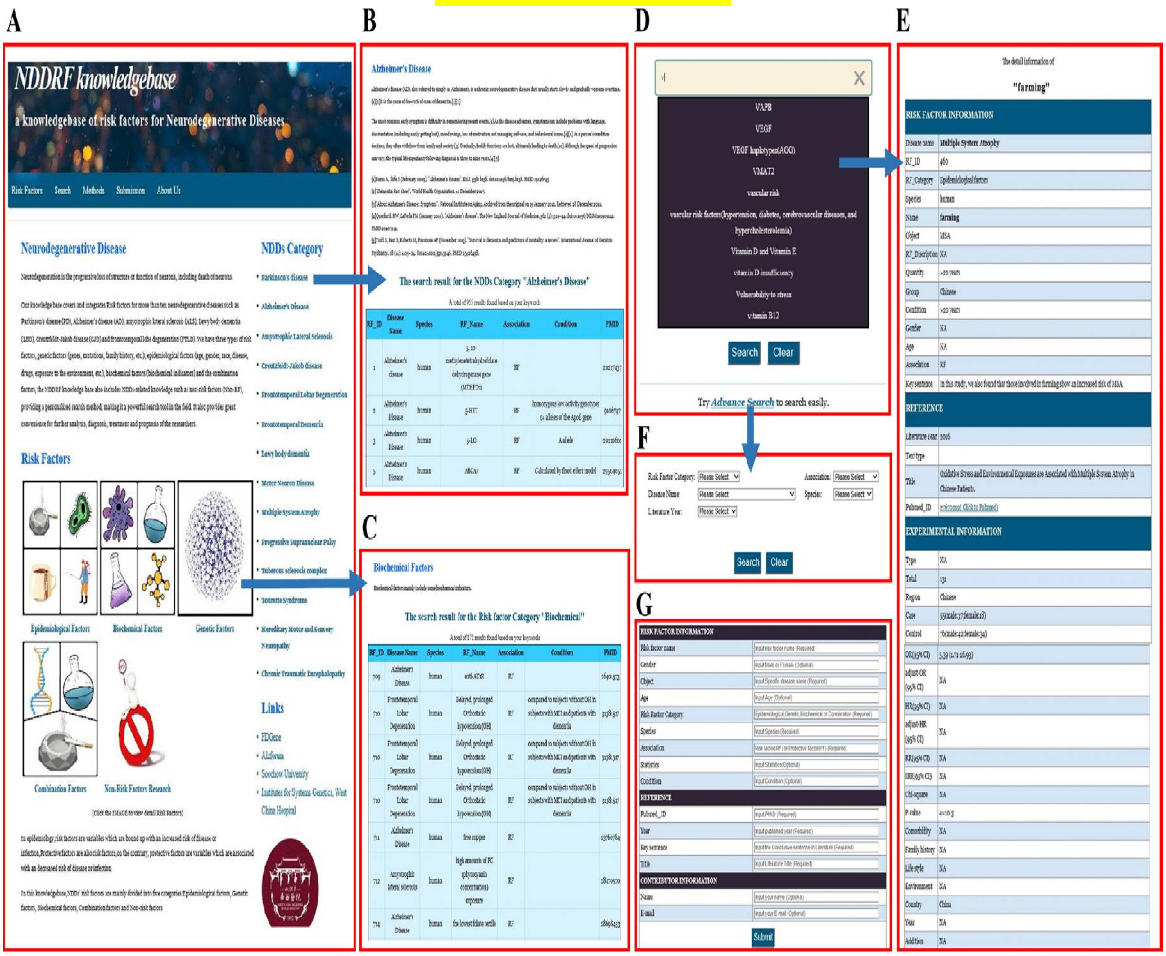
xAI.Med. | 2022-68



- ✓ DL solution is more interpretable than a DL model that predicts midline shift directly from the image

xAI.Med. | 2022-69

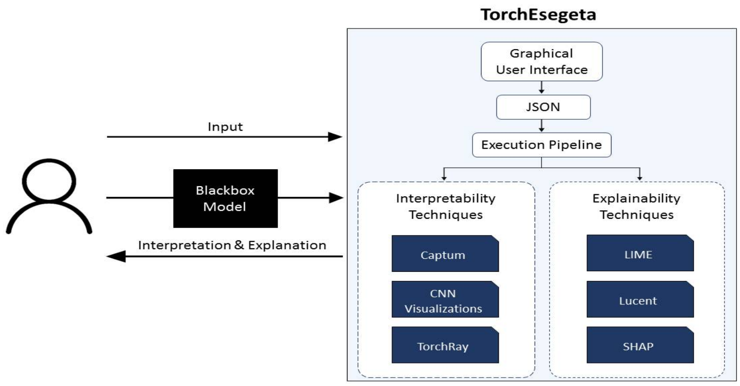
Web interface of NDDRF



(A) Risk factor page; (B) Categories of NDD; (C) Categories of risk factors; (D) Search page; (E) Details of risk factor; (F) Advanced search; (G) Submission page

xAI.Med. | 2022-211

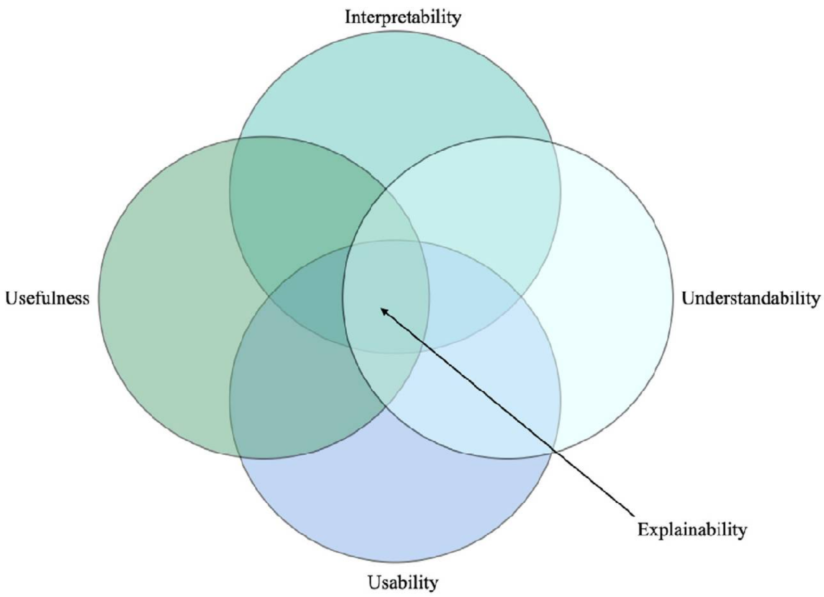
TorchEsegeta pipeline architecture



xAI

Explainability, Interpretability, Accuracy

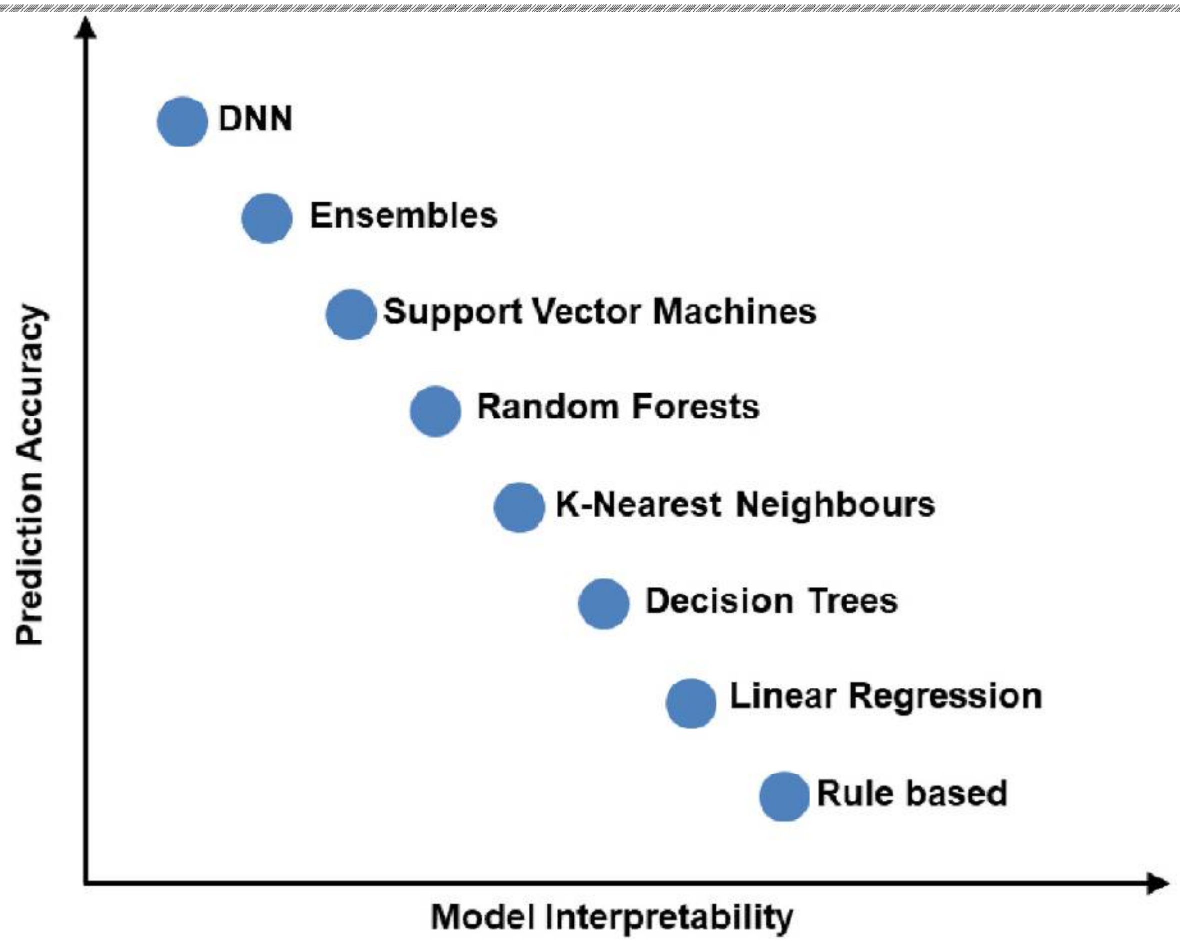
xAI.Med.		2022-17



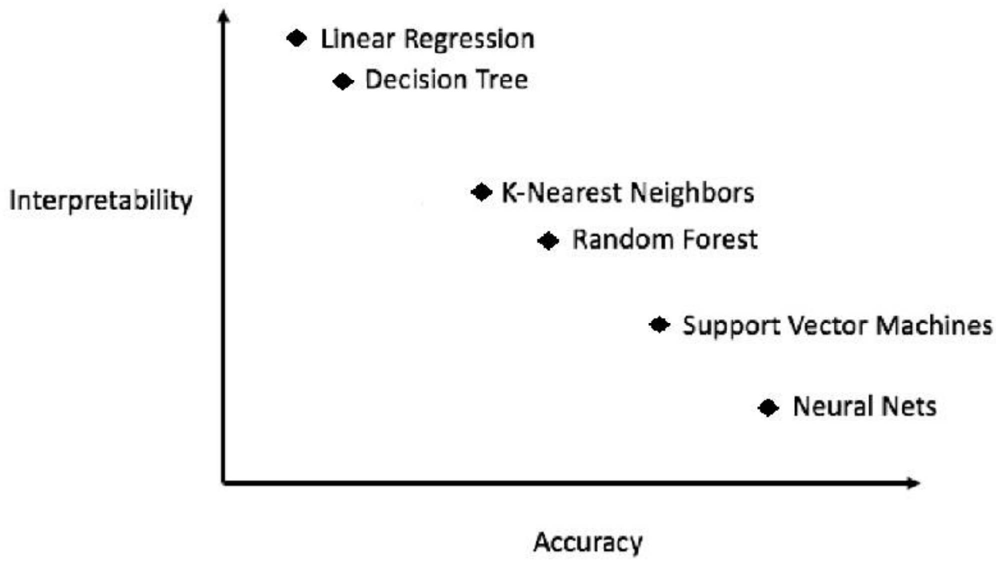
Venn diagram of explainability as
intersection of usability, usefulness, interpretability, and understandability

xAI.Med.		2023-14

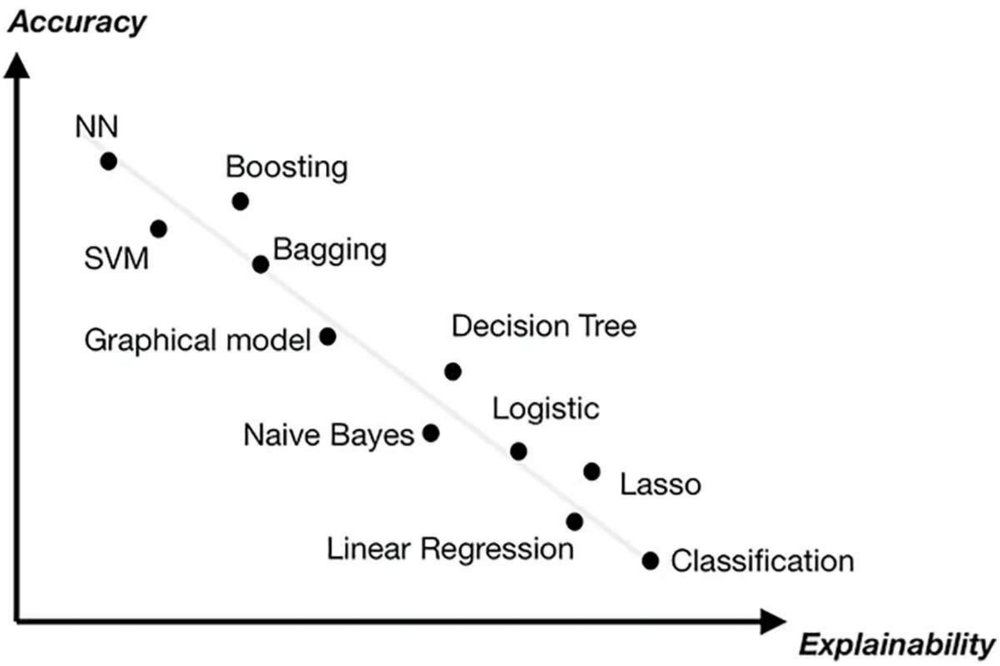
Trade-off of Model's prediction accuracy and interpretability



More accurate models are less explainable

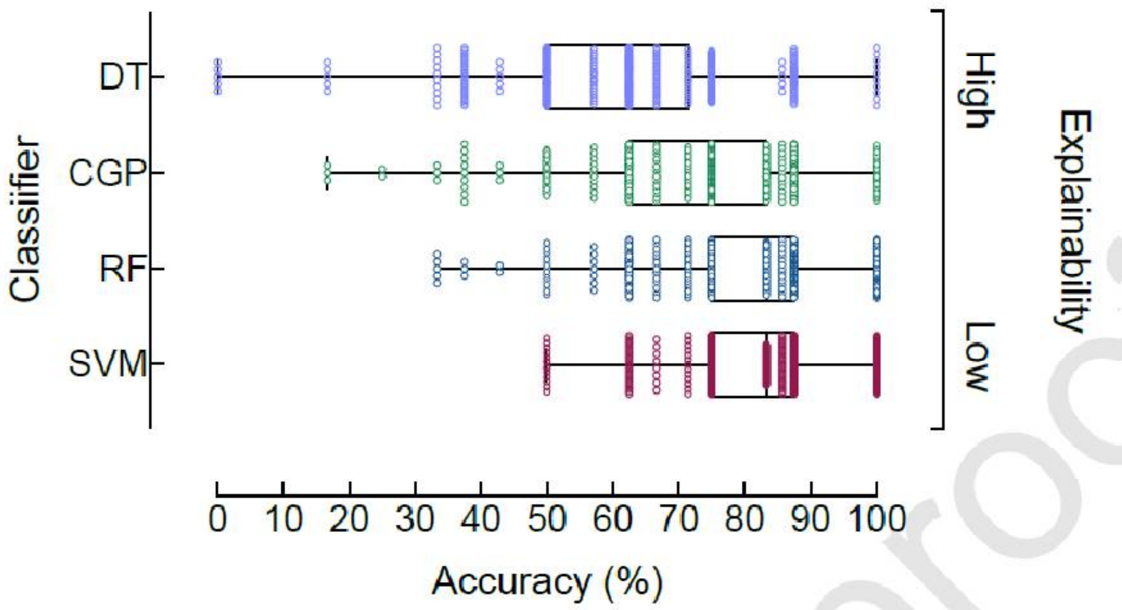


xAI.Med.		2022-205



xAI.Med.		2021-28

Trade-off between interpretability and accuracy for classifiers
On pahaw dataset

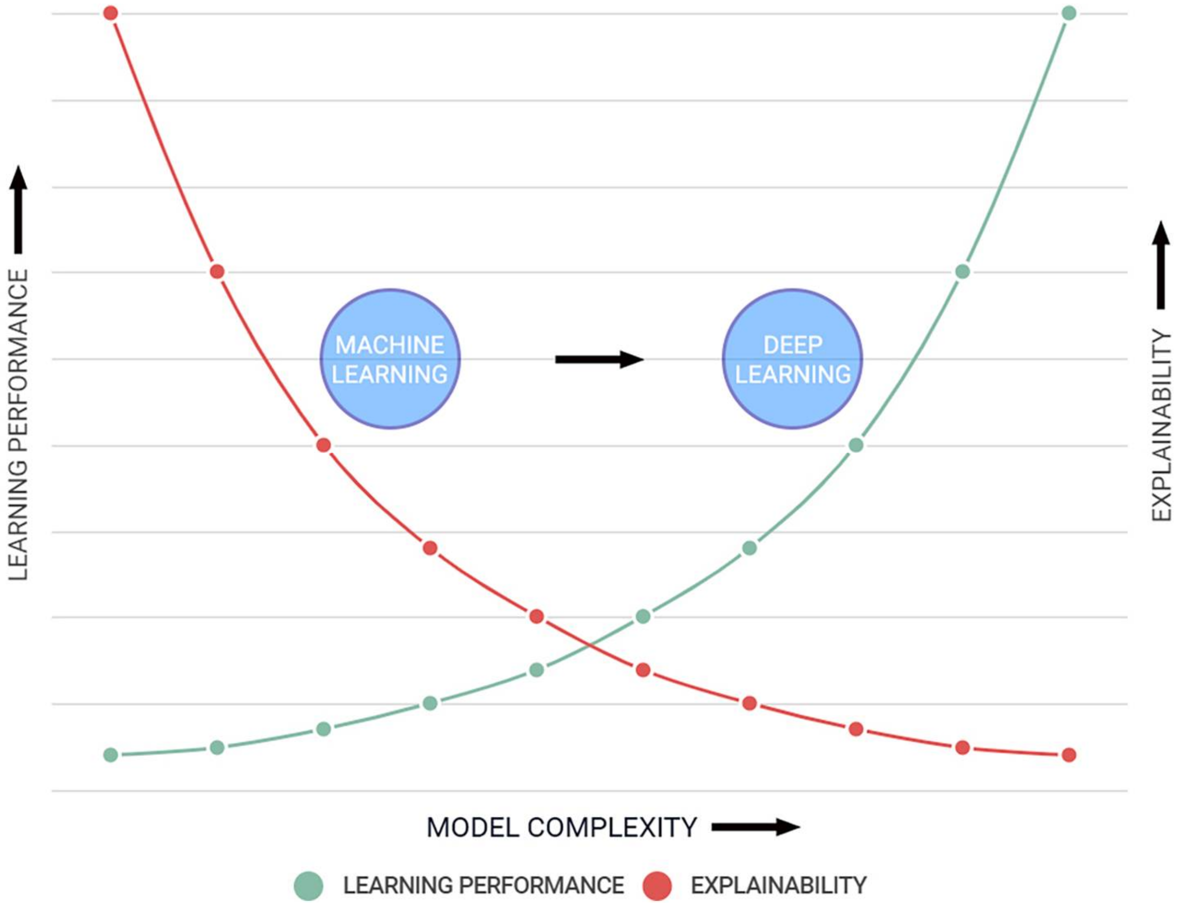


- Points :single values for the accuracy for all the runs
- Boxes: mean and the standard deviations

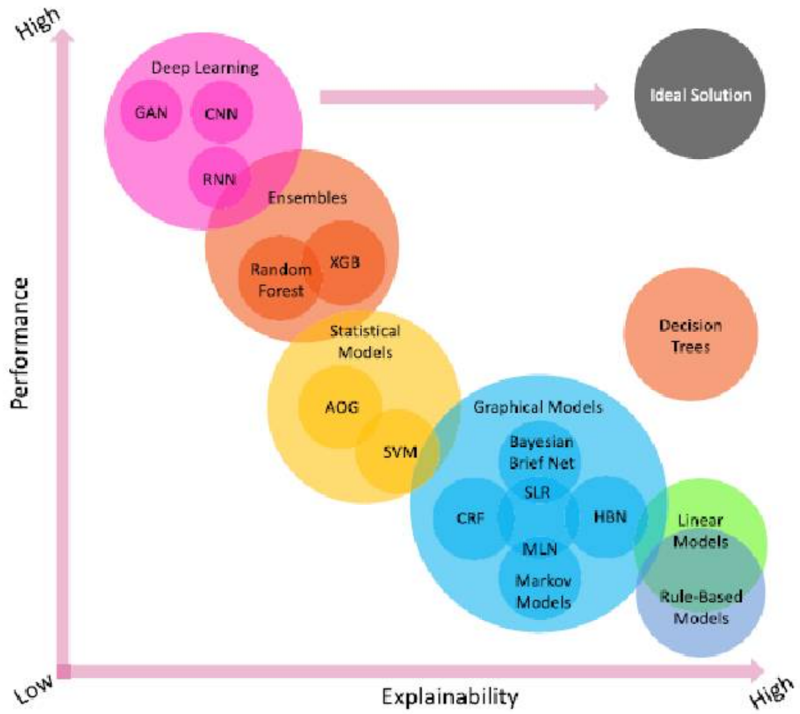
xAI.Med.		2021-110

Learning performance and explainability of AI system

as a function of model complexity



xAI.Med.		2022



HBN: Hierarchical Bayesian Networks; SLR: Simple Linear Regression; CRF: Conditional Random Fields; MLN: Markov Logic Network; SVM: Support Vector Machine; AOG: Stochastic And-Or-Graphs; XGB: XGBoost; CNN: Convolutional Neural Network; RNN: Recurrent Neural Network; GAN: Generative dversarial Network

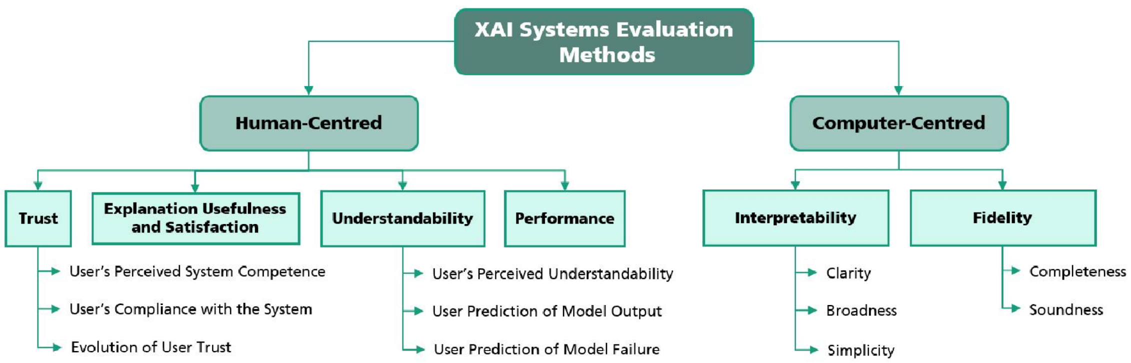
Linear models, rule-based models and decision trees

- ! More transparent, but lower performance

Complex models e.g., deep learning and ensembles

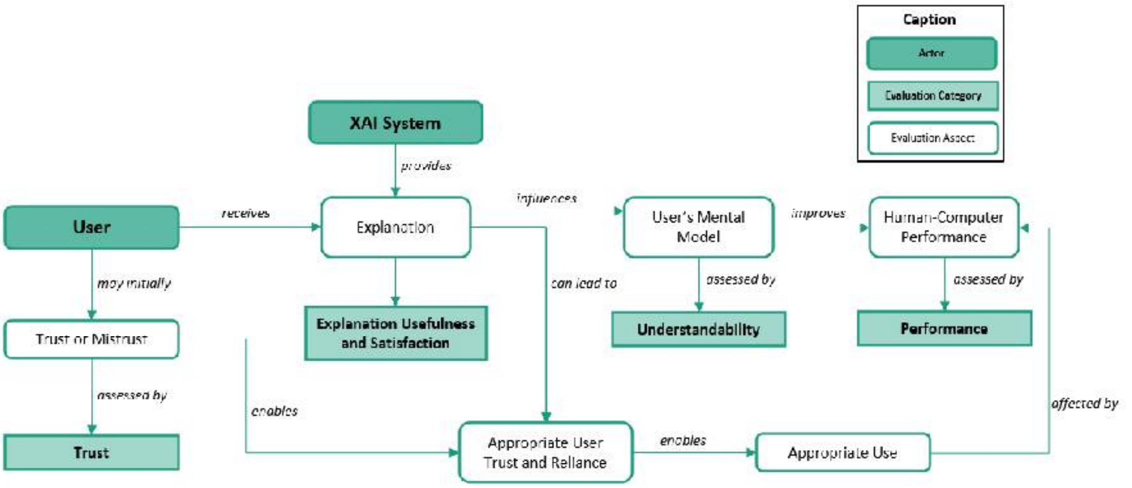
- ! Higher performance while lessexplainability

Taxonomy for XAI systems' evaluation methods



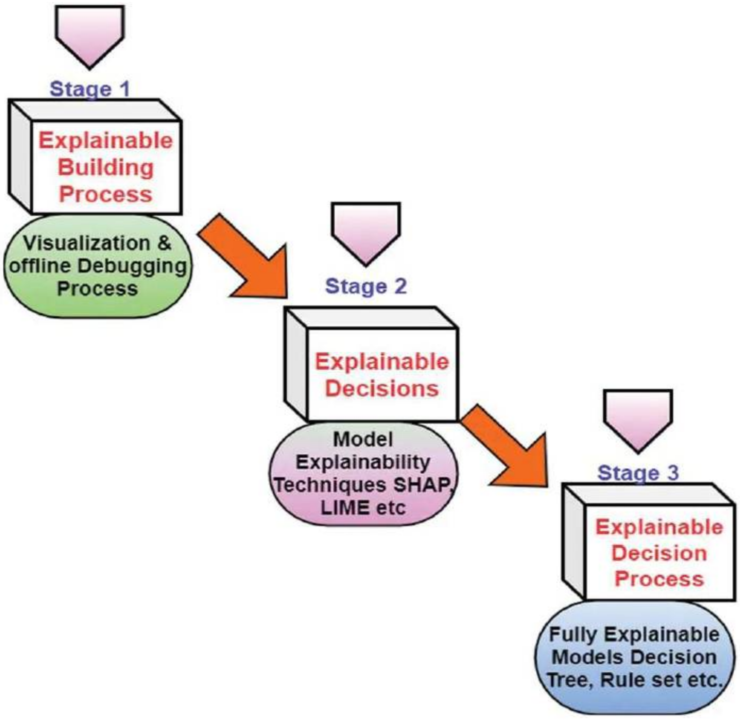
xAI.Med. 2022-210

Human-centred evaluation process



xAI.Med.		2022-205

xAI



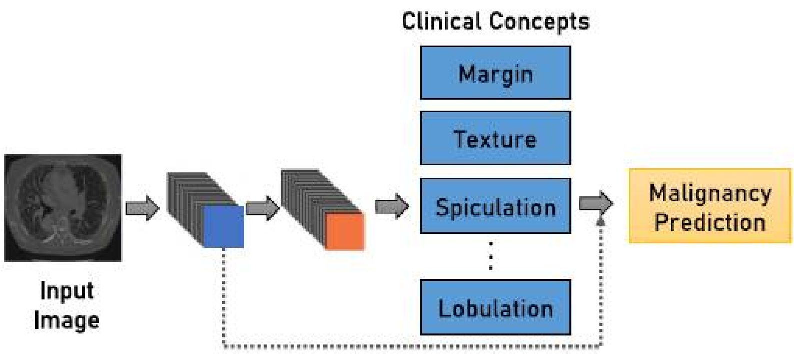
xAI.Med.		2022-62

Normal machine or deep learning procedure + explainable surrogate modules

\+ Achieves a level of more transparent and trustworthy model

xAI.Med.	2022-68

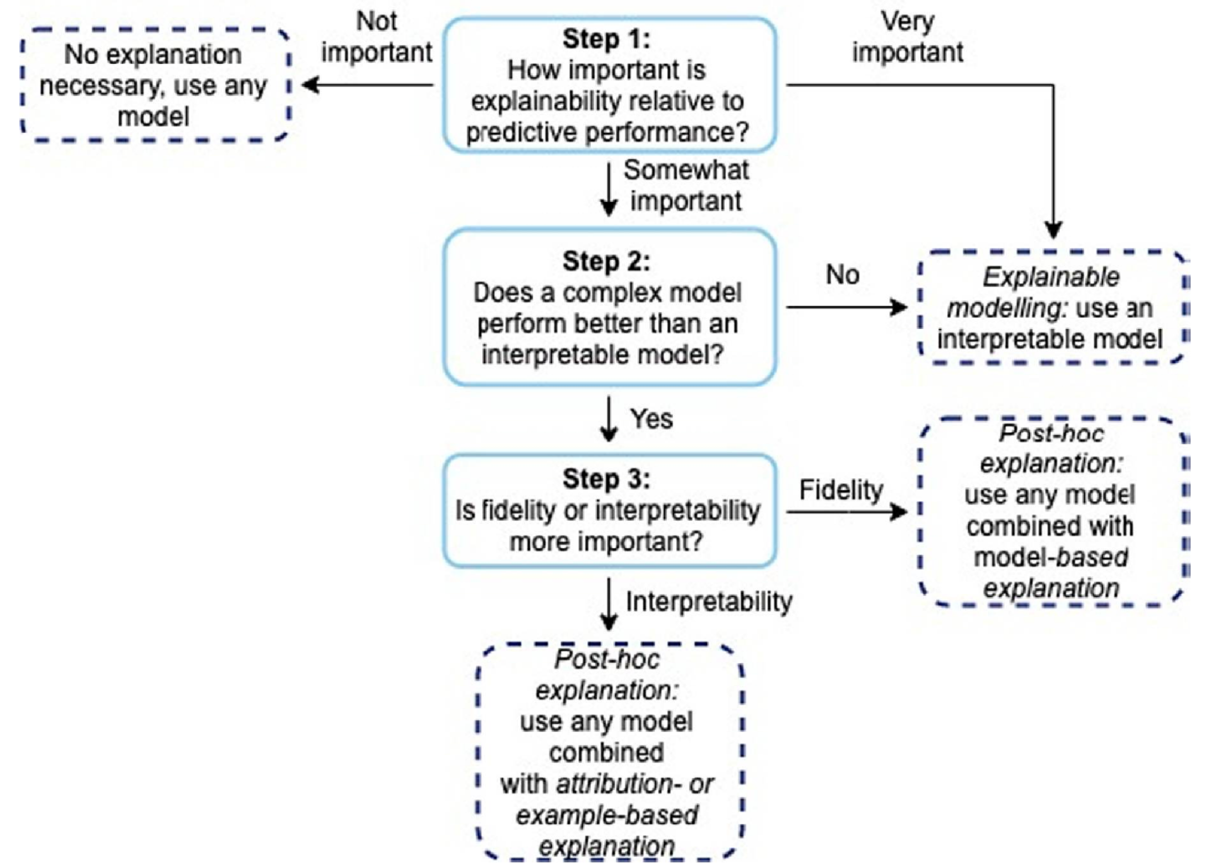
Clinical concepts + deep features



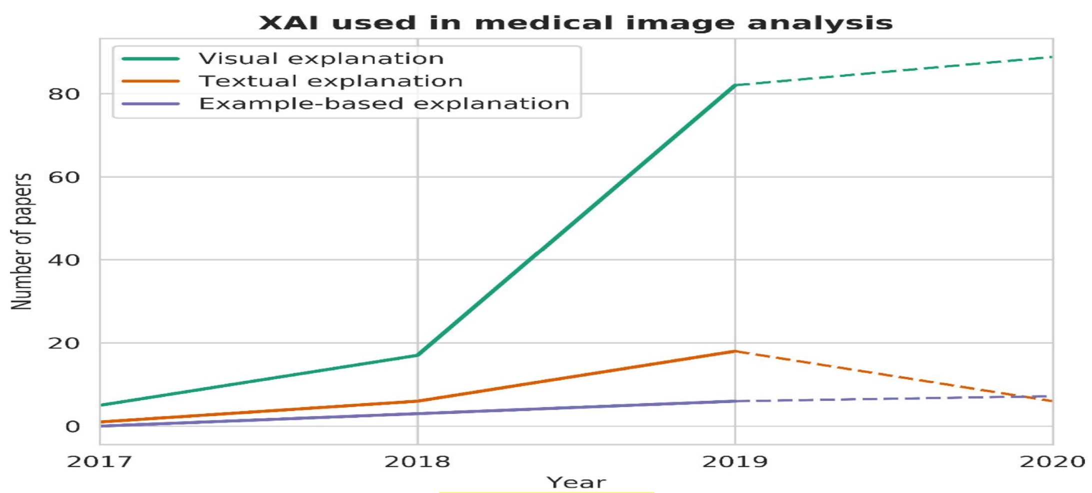
Concept Learning Models

xAI.Med.	2021-19

xAI—Need, Fancy or Essential?



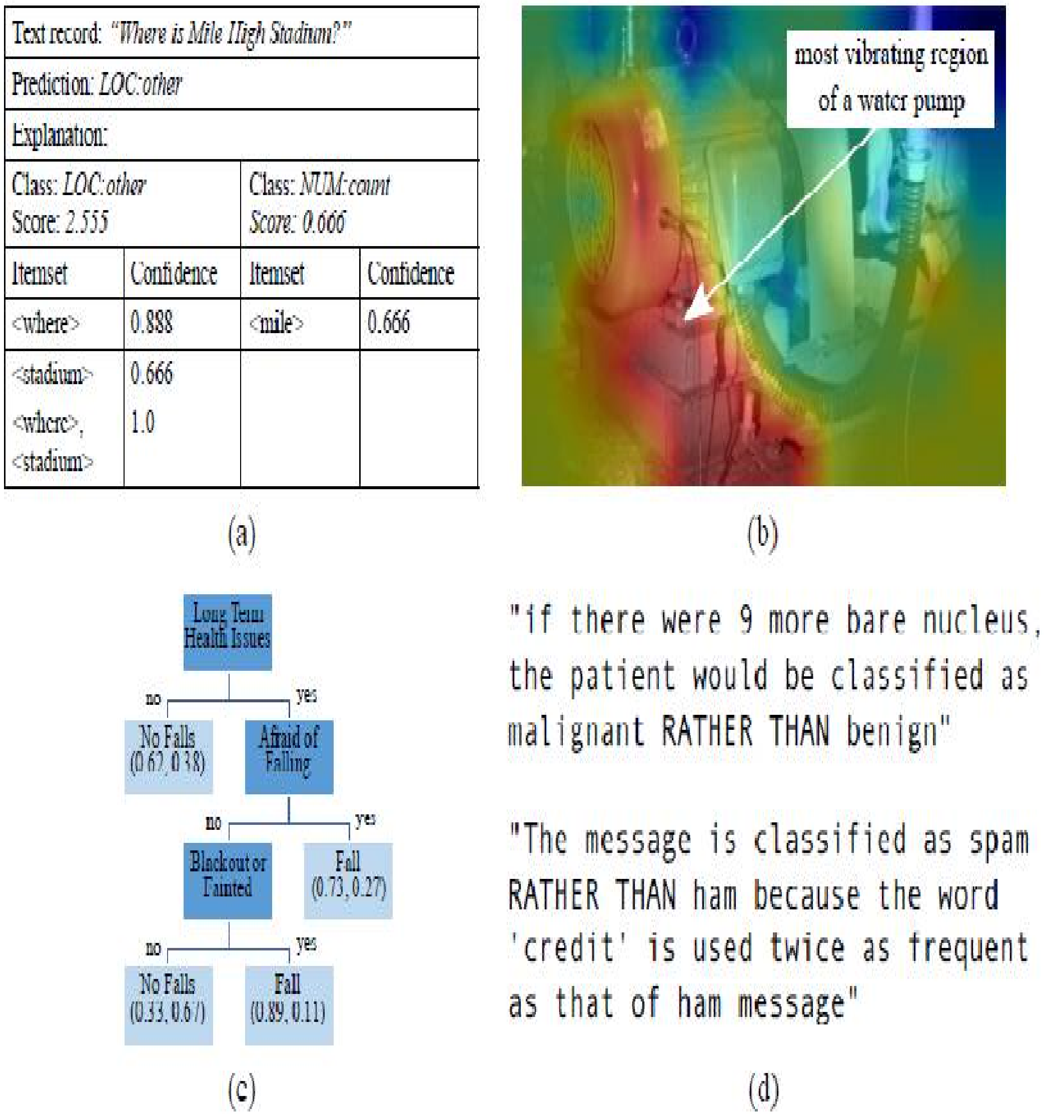
xAI.Med.		2023-57



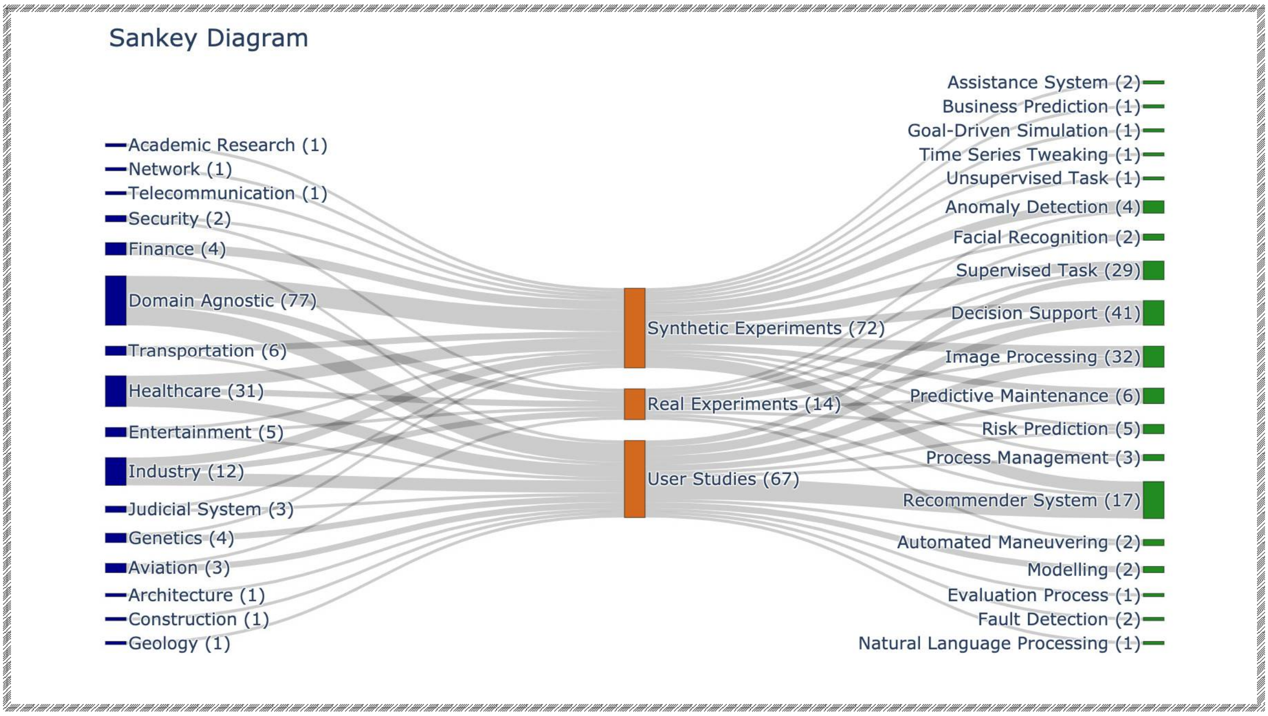
Visual explanation

xAI.Med.		2022-204

Different forms of explanations



(a) Numeric explanation of remaining life estimation in industry appliances
(b) Visual explanation for fault diagnosis of industrial equipment
(c) Example of rule-based explanation in the form of a tree
(d) Explanation text generated with GRACE

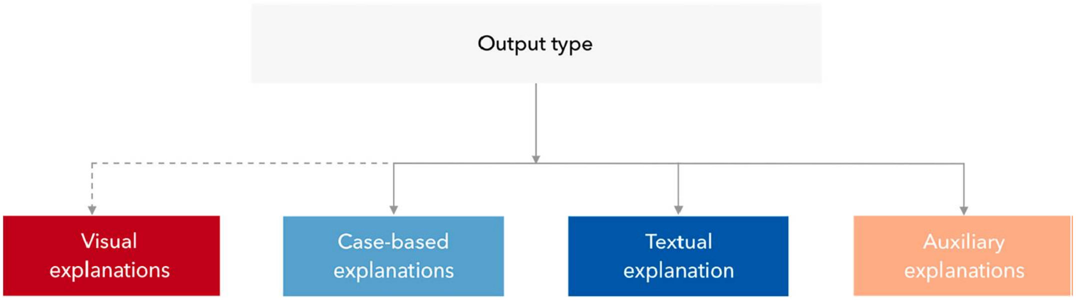


Fix AIM

Fig Imag Tab Scripts (Fits) in eXpl AI

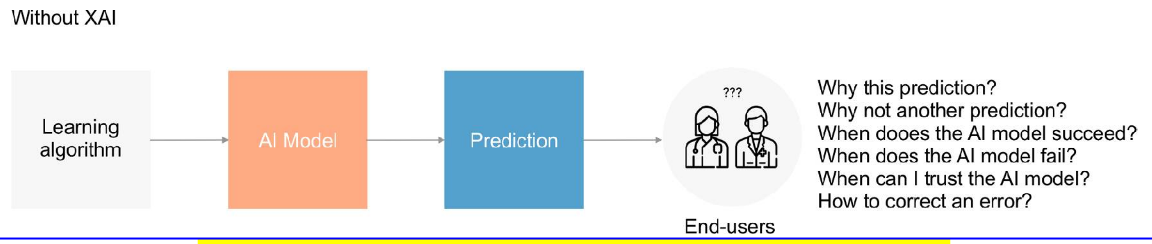
xAI.Med. | 2023-03

Distinction of explanatory approaches by resulting output presentation form



Non-visual XAI approaches encompass auxiliary, case-based, and textual explanations

xAI.Med.	Case Study	2023-22

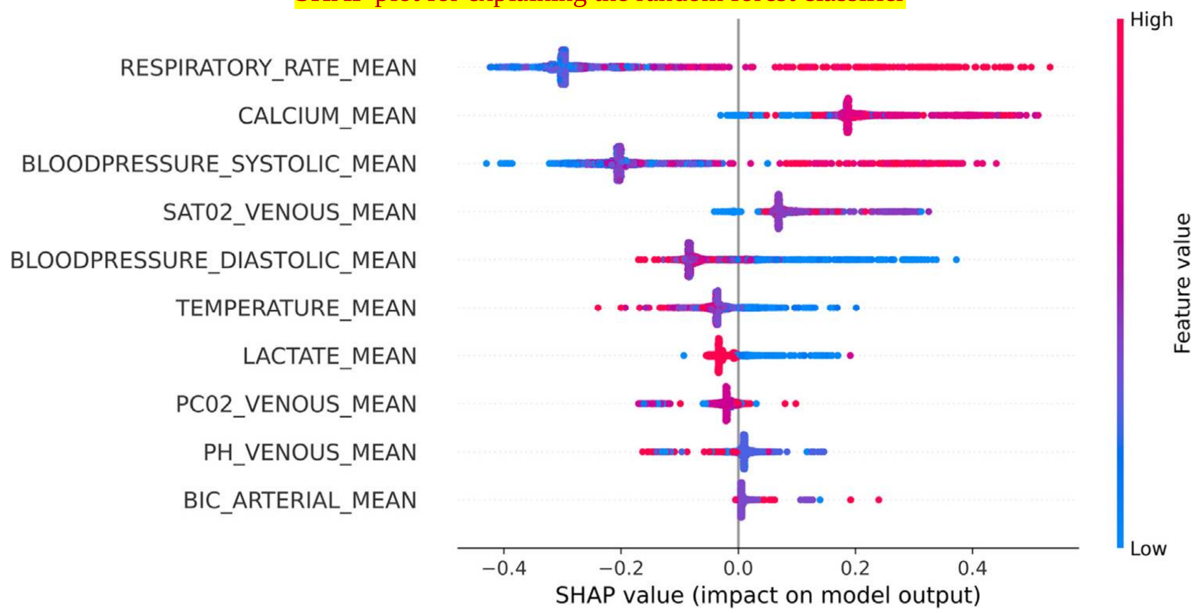


Interaction between AI models and end-users without use of XAI

SHAP plot

xAI.Med.		2023-17

SHAP plot for explaining the random forest classifier

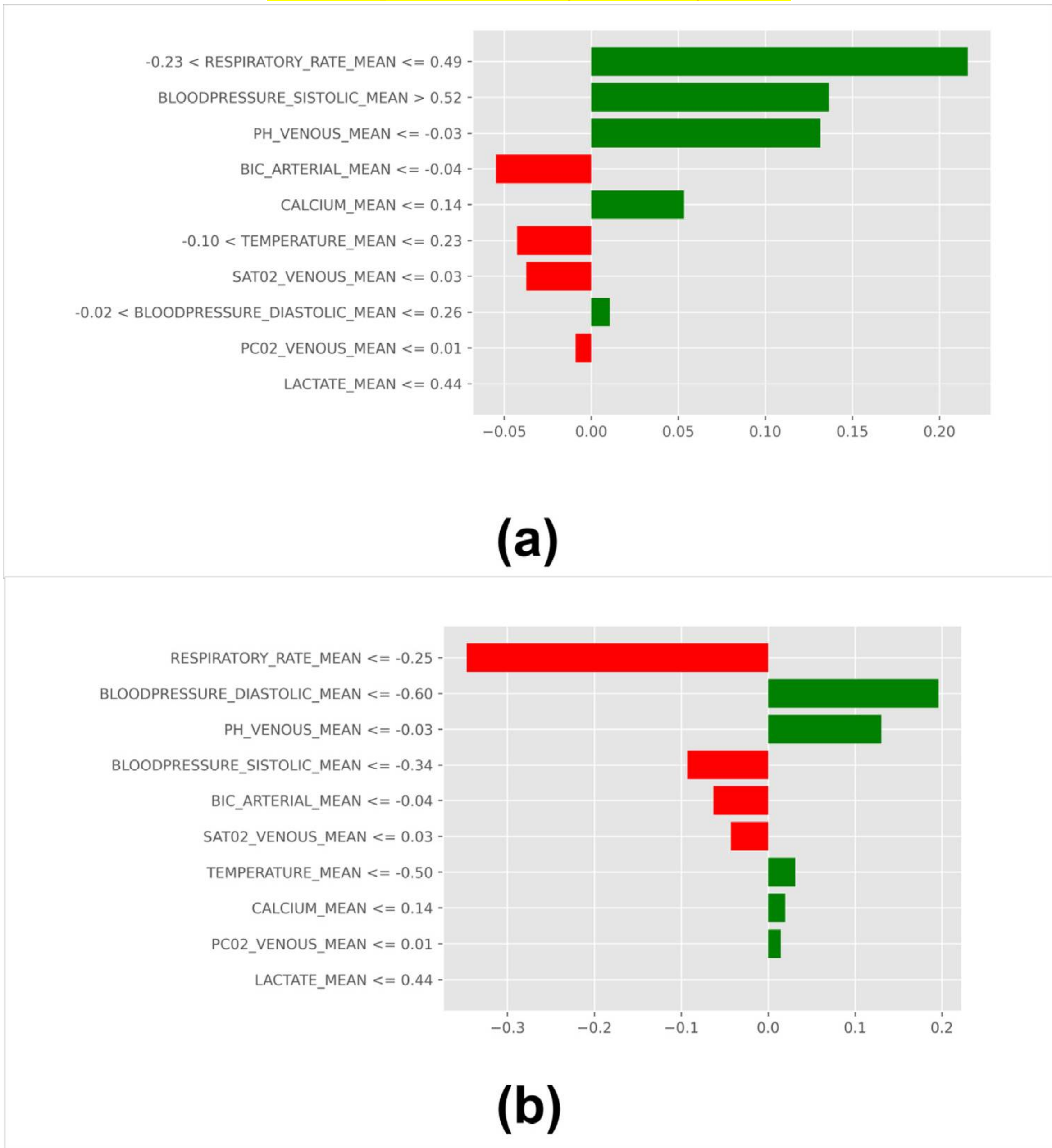


- o Respiratory Rate mean has the most significant contributor to the output.
- o Higher values of features such as Respiratory Rate mean, Calcium mean, and Systolic Blood pressure value had a positive correlation with the SHAP values.
    - o This positive relation indicates that a higher value of these features could increase the patient's probability of being classified as severe (in need of an ICU)
- o Ground truth: These findings coincide with the research conducted regarding COVID-19 severity

Local Interpretable Model-Agonist

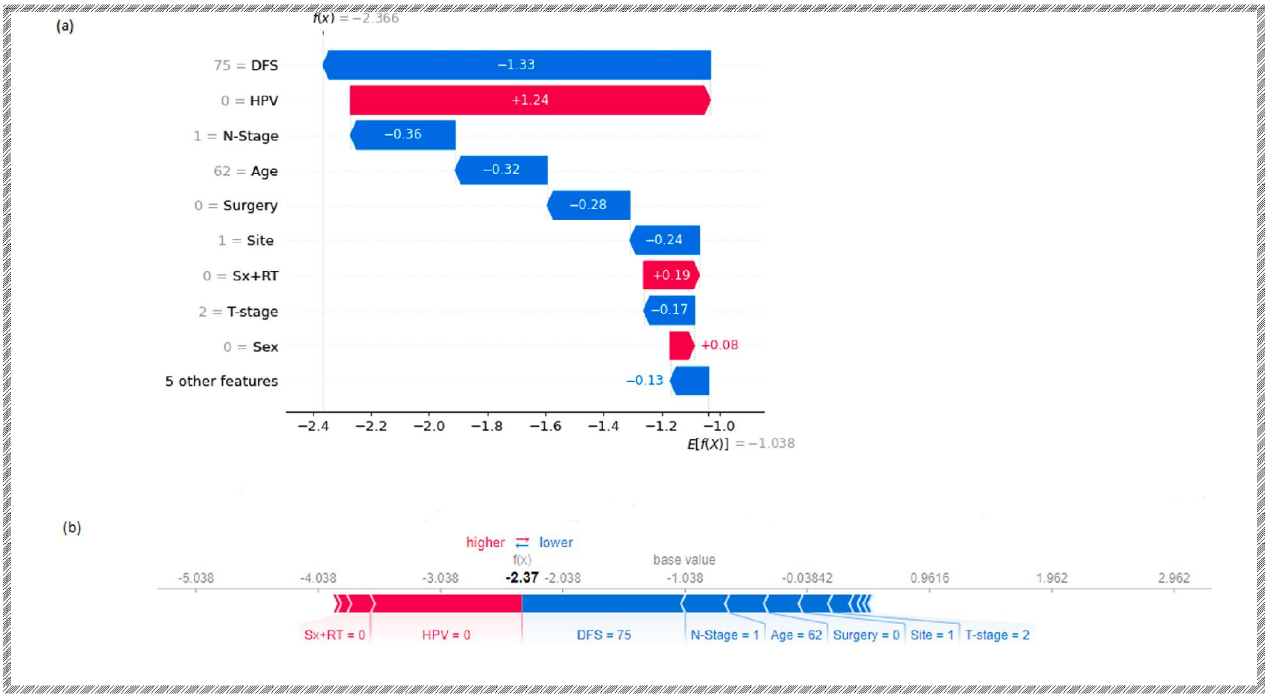
xAI.Med.		2023-17

Local Interpretable Model-Agonist for Light GBM

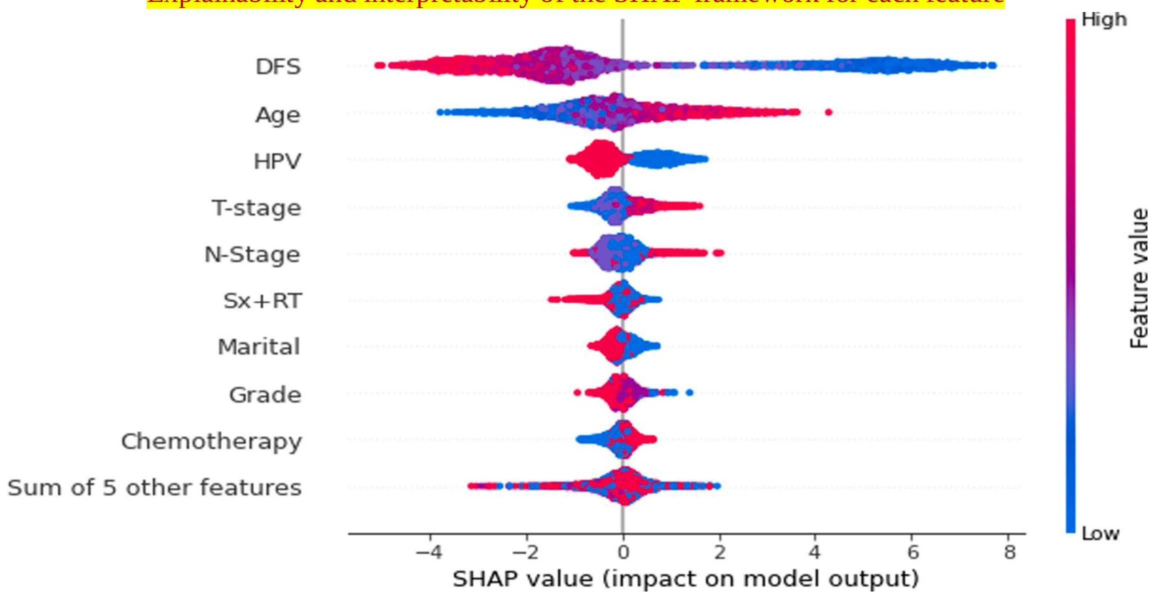


xAI.Med.		2022-12

Shapley Additive Explanation (SHAP) framework for individual predictions



Explainability and interpretability of the SHAP framework for each feature



LIME Plot

Local Interpretable Model Agnostic Explanations

xAI.Med.		2022-12

LIME framework for individual predictions

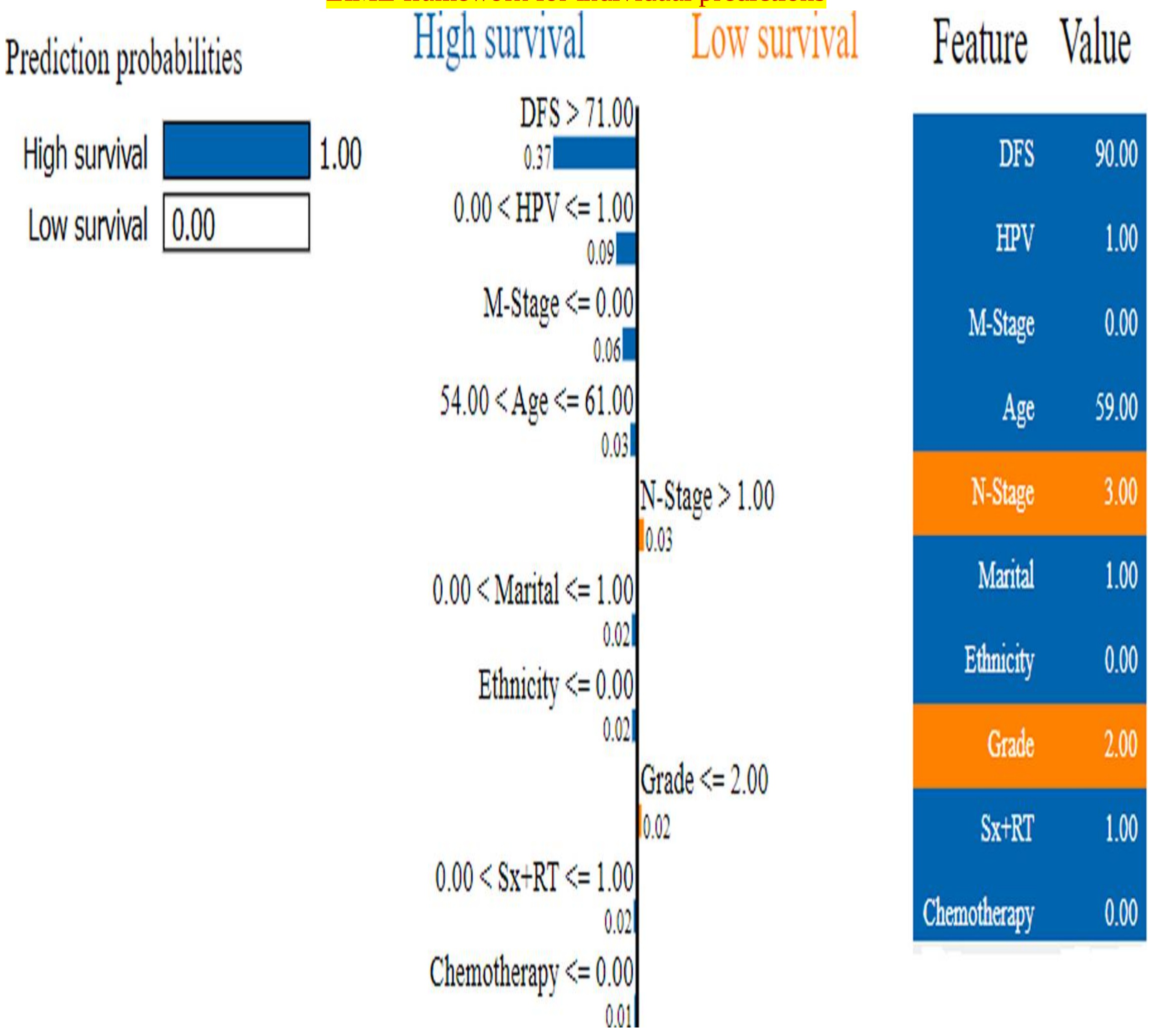
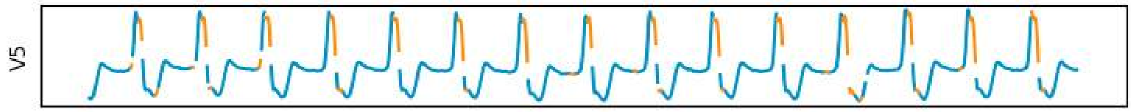


Fig. 4. The Local Interpretable Model Agnostic Explanations (

xAI.Med.		2022-24

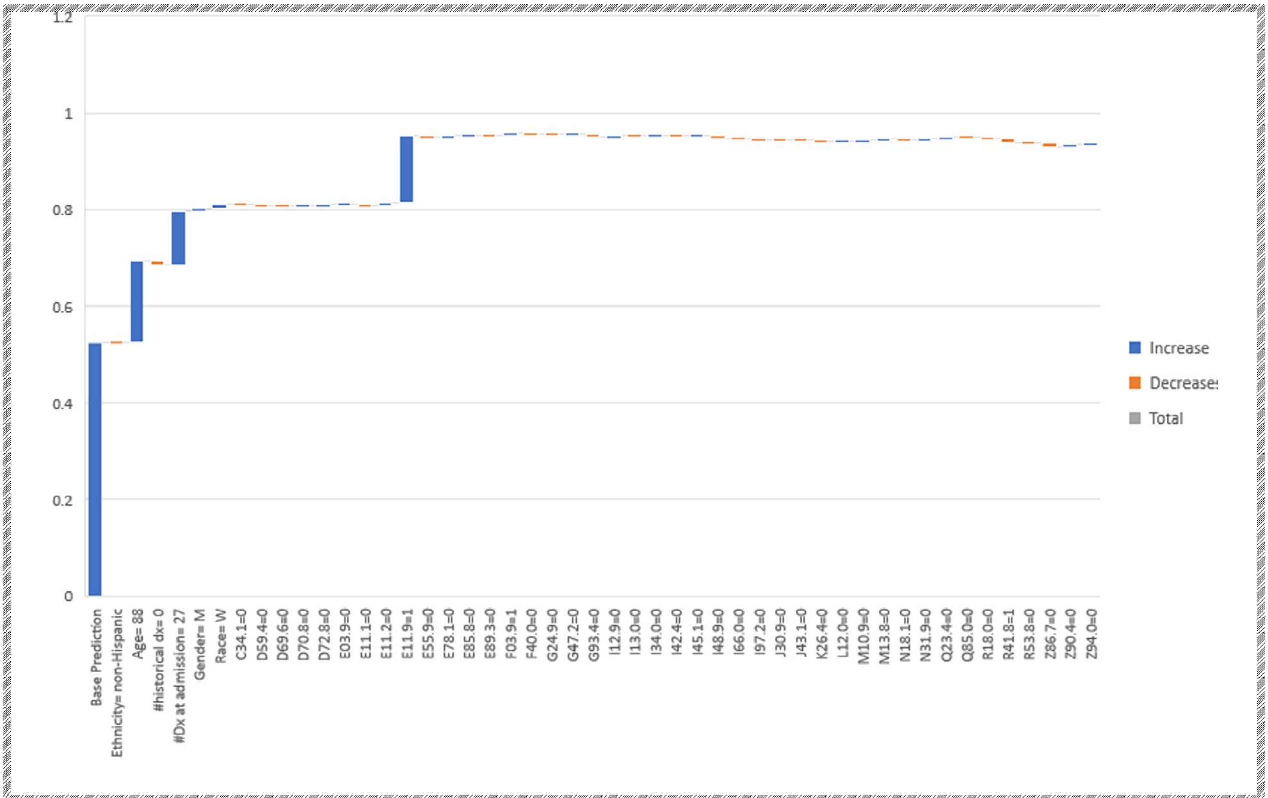
Result of interpretation by the SHAP method for the ResNet neural network.
Diagnosis – (Conduction Disturbance) CD



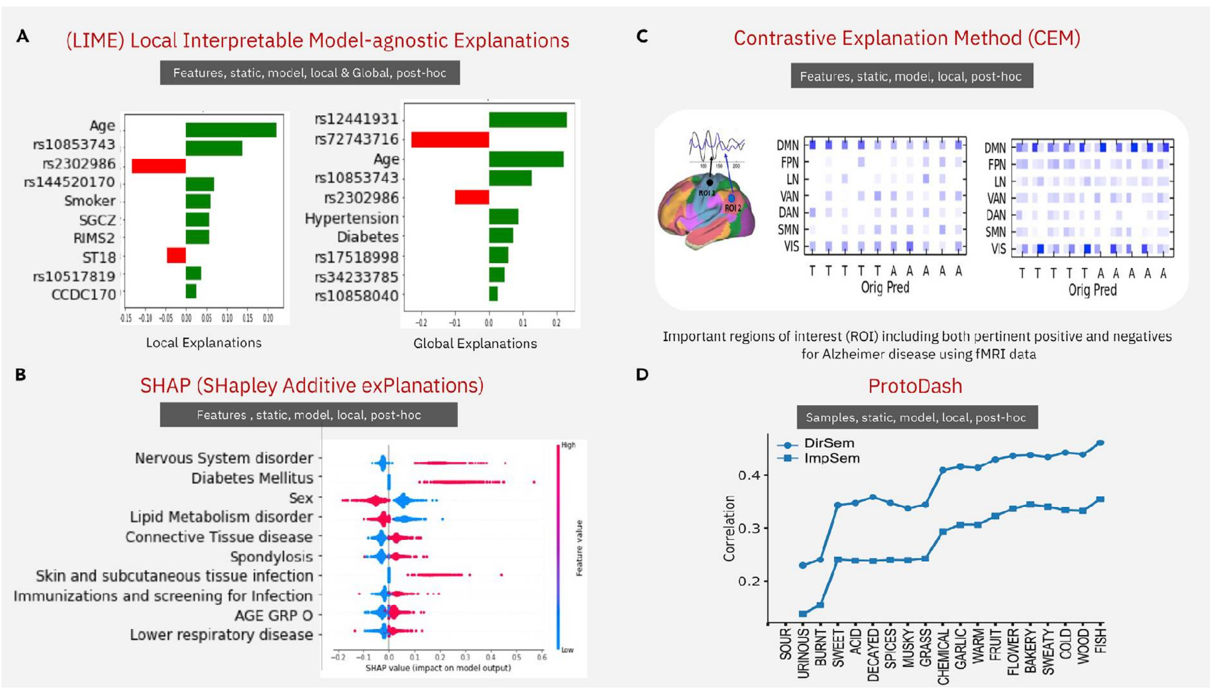
✓ Areas that make the greatest contribution to the forecast are orange

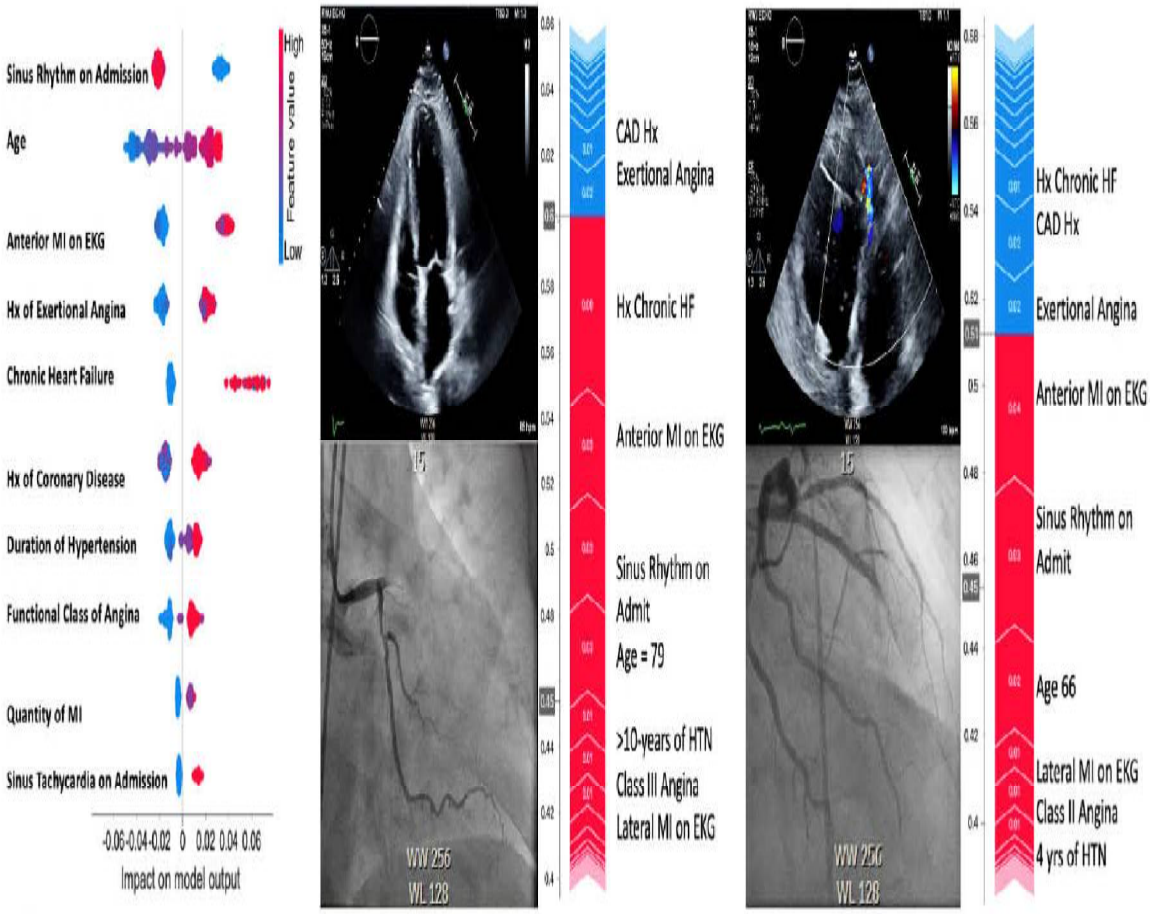
xAI.Med.		2022-27

SHAP scores for an individual patient (feature values are given on the horizontal axis).



Applications of four popular XAI methods





A. Shapley Plot for 20% Internal Validation

B. Cardiogenic Shock Case: Echo, Coronary Angiography (100% prox LAD), Individual Shapley Plot

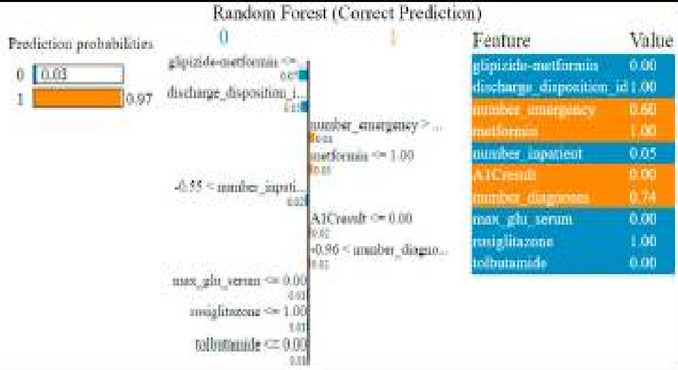
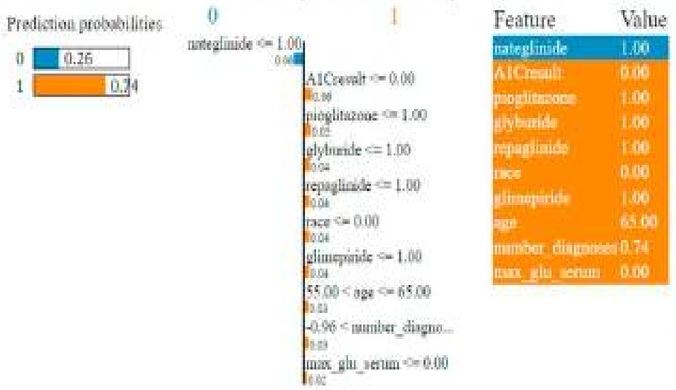
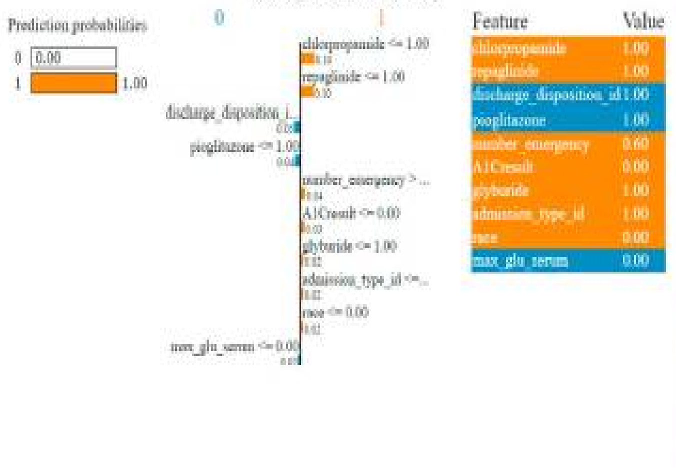
C. VSD Case: Echo, Coronary Angiography (100% prox LAD), and Individual Shapley Plot

Red: Favoring Model Prediction of MI Complication
Blue: Opposing Model Prediction of MI Complication

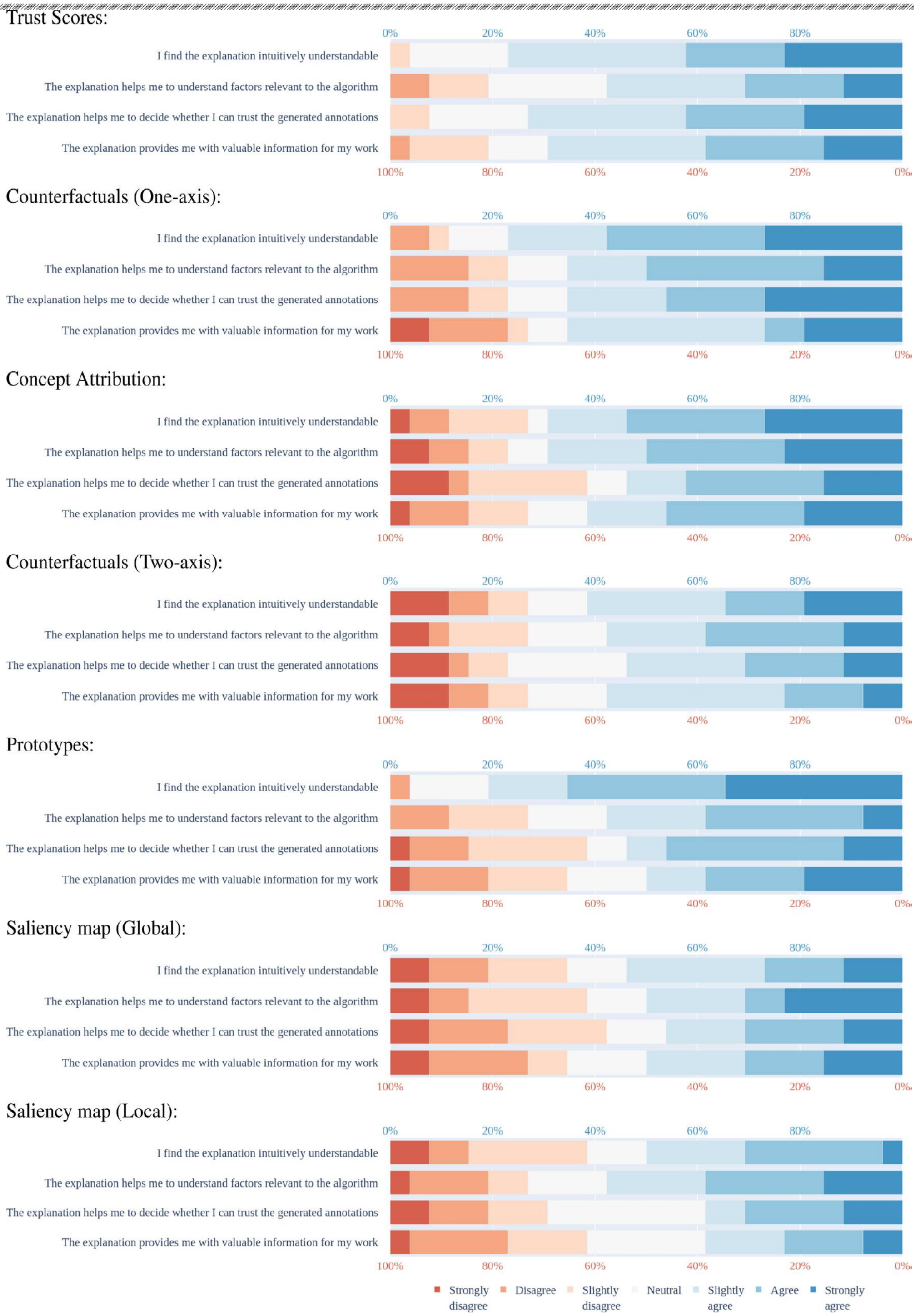
Figure A :Key variables enabling prediction of complications including age, anterior MI, exertional angina, chronic heart failure, prior CAD, prior MI, andsinus tachycardia.
Figure B and C:External predictions in two patients with MI complications (cardiogenic shock and VSD, respectively) where relative features favoring (red) and those opposing the model prediction (blue) are delineated.

✓ **Conclusion:** ML and explainable AI can be used to integrate ECG and clinical characteristics to predict complications following MI

xAI.Med.		2022-40
LIME Visual Explainer for Patient 1 (Readmitted < 30 days)		

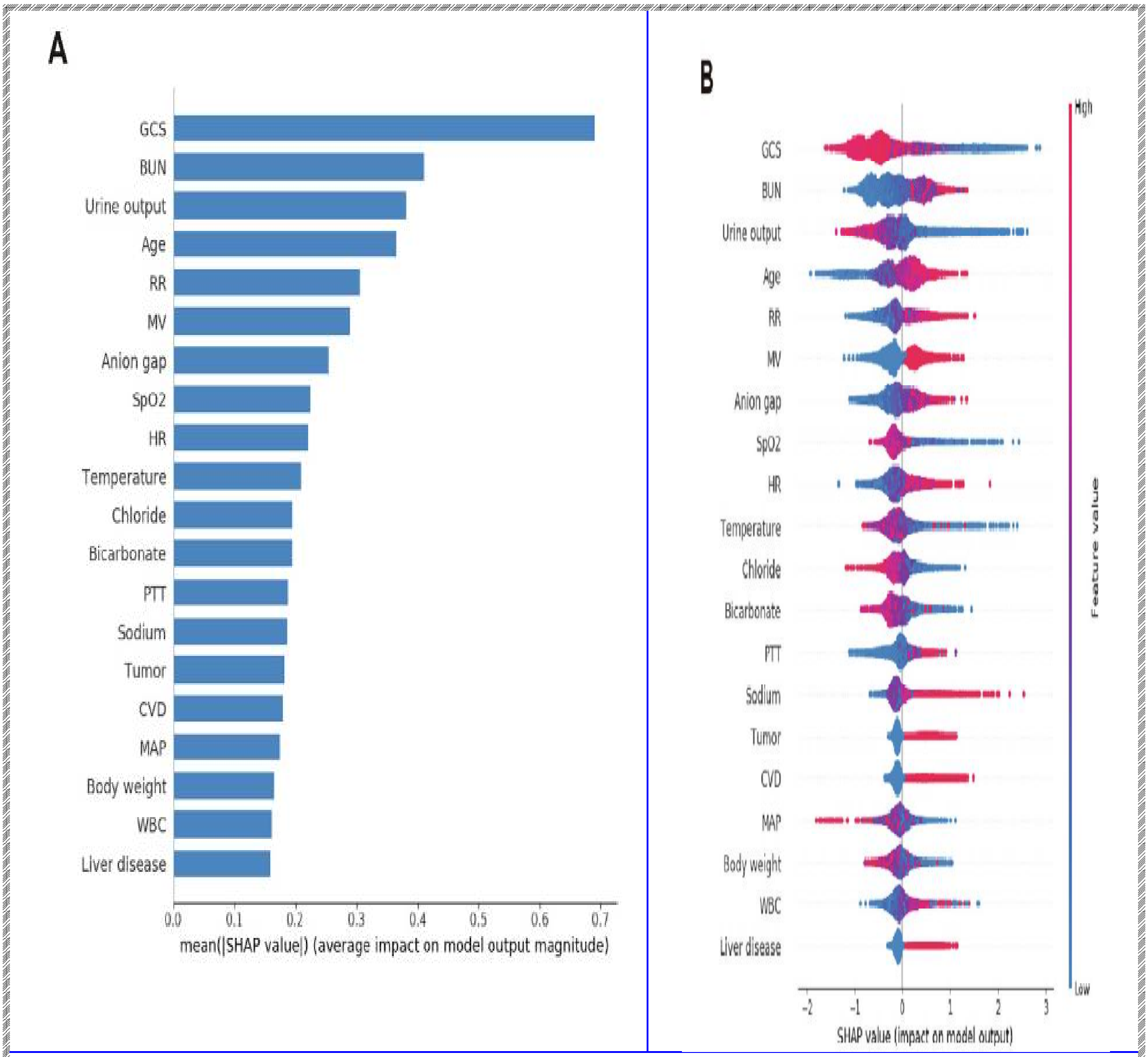
Patient 1: (Ground Truth: Class 1 = Readmitted < 30 Days)	
Random Forest (Correct Prediction) 	For RF, the most important features that led it to predict that this diabetic patient is for readmission in < 30 days are: • high number of emergency visits (number_emergency=0.60), • not taking metformin (metformin=1), • not monitoring his HbA1c (A1Cresult=0), and • high number of diagnoses (number diagnoses=0.74). For the domain expert, not taking diabetes medications when a confirmed diabetic, frequent visits to the emergency room, and high number of diagnosis constitute a high risk for readmission in < 30 days which is in agreement with RF's logic for this patient.
AdaBoost (Correct Prediction) 	For AdaBoost, this patient is a diabetic but • is not taking any diabetes medications (pioglitazone=1, glyburide=1, repaglinide=1, glimepiride=1), • is elderly (age=65), • has many other diseases (number_diagnoses =0.74), • is not monitoring his HbA1C (A1CResult=0), and • has no glucose serum on file (max_glucose_serum =0). Hence for AdaBoost, these are grounds to predict that this patient is a candidate for readmission in < 30 days. For the domain expert, not taking diabetes medications when you are a diabetic plus presence of many other diseases and old age are possible grounds for readmission in < 30 days. This is in agreement with AdaBoost reasoning for this patient
KNN (Correct Prediction) 	For K-NN, this patient is diabetic but • is not taking any diabetes medications (chlorpropamide=1, repaglinide=1, glyburide=1), • is not monitoring his HbA1C (A1Cresult=0) and glucose serum (max_glu_serum=0), • has high number of emergency visits (number_emergency=0.6), • this particular admission is an emergency admission (admission_id=1) Hence, for K-NN, these are grounds to predict that this patient is a candidate for readmission in < 30 days. For a domain expert, not taking diabetes medications when a confirmed diabetic, has a history of frequent emergency admissions and with this admission as another emergency admission are grounds to predict possible readmission in < 30 days. This is in agreement with the reasoning used by K-NN for this particular patient.

Distribution of responses from 25 admitted online questionnaire submissions



xAI.Med.		2022-44

SHAP summary plot for the top 20 clinical features

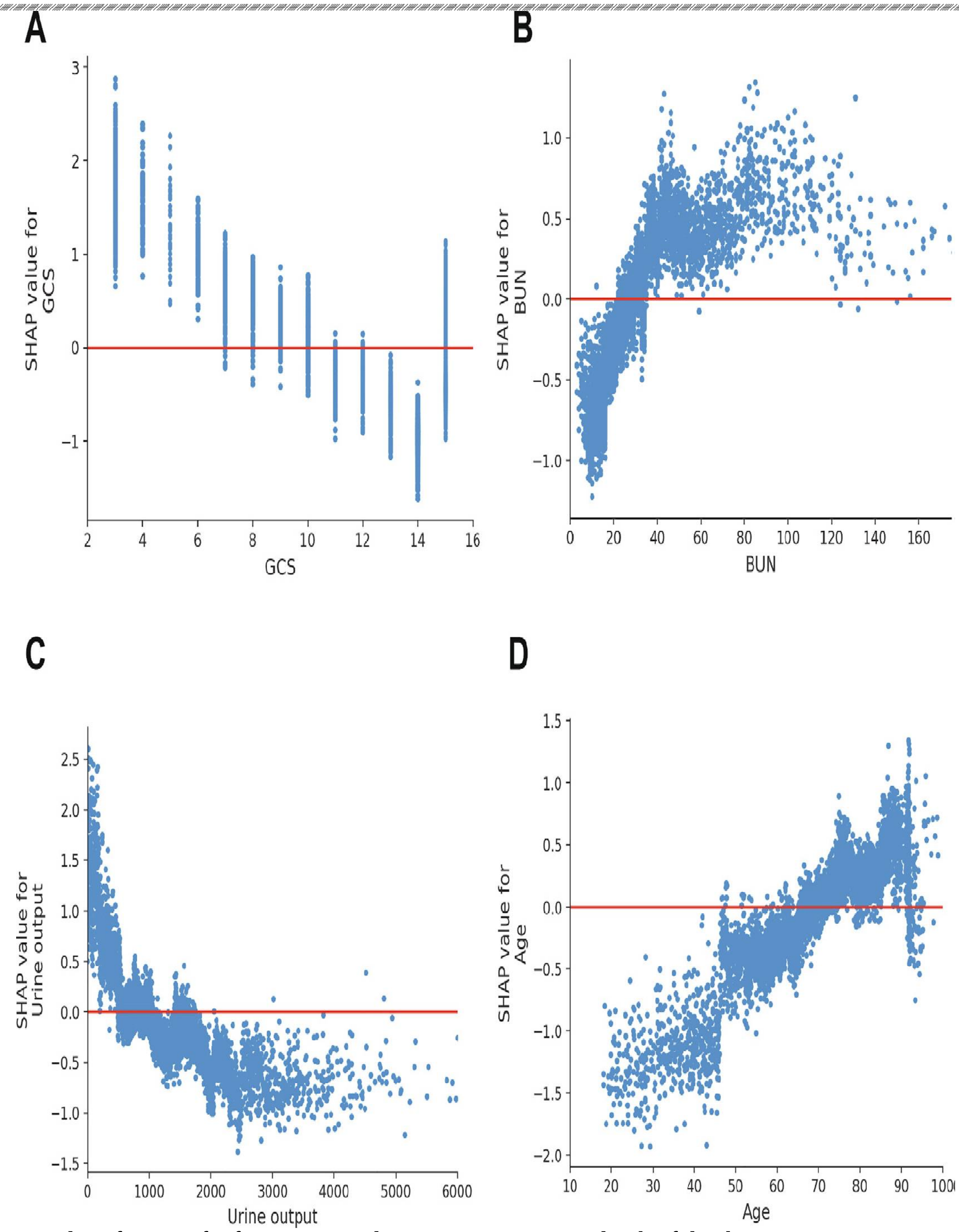


Features :GCS, Glasgow Coma Scale; BUN, blood urea nitrogen; RR, respiratory rate; MV, mechanical ventilation; HR, heart rate; PTT, partial thromboplastin time; CVD, cerebrovascular disease; MAP, mean arterial pressure; WBC, white blood cell.

xAI.Med.		2022-44

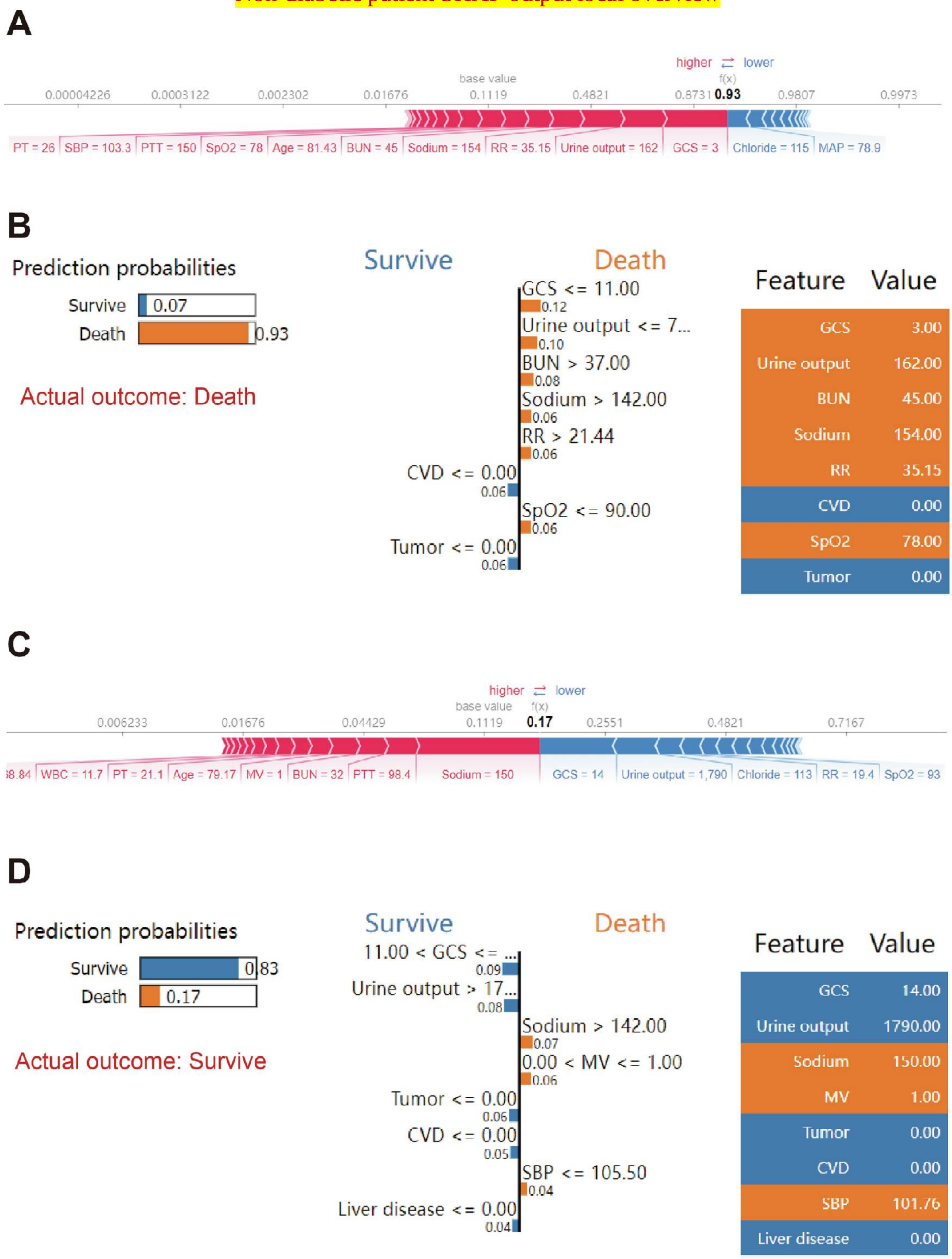
SHAP dependence plot for clinical features

(A) GCS; (B) BUN; (C) Urine output; (D) Age



IFSHAP values for specific featuresexceed zero, **THEN**increased risk of death

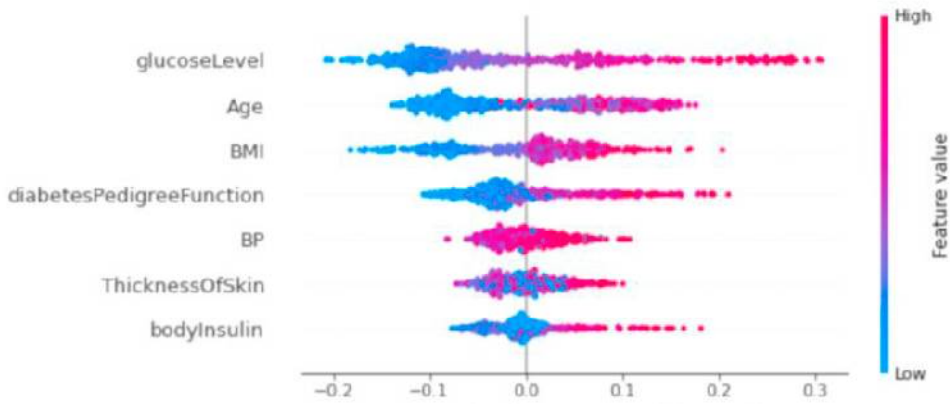
Non-diabetic patient SHAP output local overview



SHAP Global Explanation and Model summary

```
[109] # SHAP global explanation
      shap_model = shap.TreeExplainer(rf)
      shapely = shap_model.shap_values(X_train)

shap.summary_plot(shapely[1], X_train, plot_type = 'dot' )
```



xAI.Med. | 2022-52

SHAP values for an individual patient