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

ISSN: 2278-1862

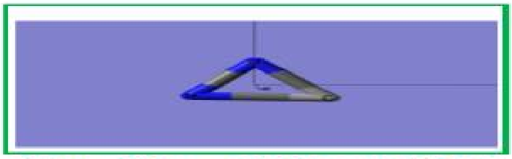
Journal of Applicable Chemistry

2023, 12 (5): 779-834

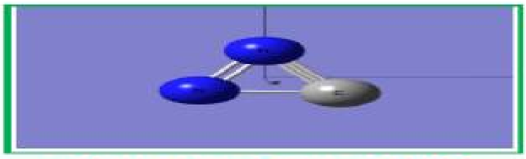
(International Peer Reviewed Journal)

New Chemistry News

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



New News of Chem (NNC)



ChemNewsNew (CNN)

CNN-57--Fit (Figure Image TableScript...)BasesPart 5. xAI(Bfit) 2022-2023 Applications

Information Source	sciencedirect.com;ACS.org ;	
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:AI and xAI play pivotal role in the outcome of collaboration of top-class (expert)researchers in science, technology, humanities and engineering.The break-throughs in Health care, environment and defense sector change world scenarios and promote hope for prosperous, peaceful, comfortable and happy human life on the lap of mother Nature.

xAI; The explainability in symbolic expert systems (likeMycin, Dendral etc.,sub set of AI-basedproducts) dates back to early 1980s, the era of Buchanan &Shortliffe.
In 2015, DARPA (Defense Advanced Research Projects Agency), USA initiatedthe disciplinexAI (e**x**plainable Artificial Intelligence)with a primary goal of enabling end users/stack-holders to better understand, trust, and effectively manage artificially intelligent intricate systemsin civilian life and

Défense operations. xAI-embedded tools/frame-works/products explain the inner process of a models, methods, procedures, data-flows and output of the processes. In 2017, a 4-year XAI research program began with multiple criteria viz. scientific consensus, medical reasoning, knowledge recall, bias, and likelihood of possible harm. The evaluation by clinicians and non-clinicians from a range of backgrounds and countries was planned. By now (year 2023) the state-of-knowledge-xAI outreached all portals of Science/engineering/ technology with a promising future of knowledgeable society.

Application Disciplines of xAI: xAI encompasses now machine learning, NNs, Deep-architectures, Deep learning and in future may be hybridized/integrated/fused or evolve into a new form in all rational endeavours from sub-atomic to astronomical material scale and wide range of energy levels. The basic disciplines with large impact of xAI are physics, chemistry and biology at macro-/ micro-/ nano-/ molecular-/ atomic levels. The applied and trans-areas of concern are medicine, Molecular/material properties, environment, synthesis, proteins etc. The results (Fit: Figure Image Table Script …. Bases) of typical case studies during 2022Jan to 2023June are documented.

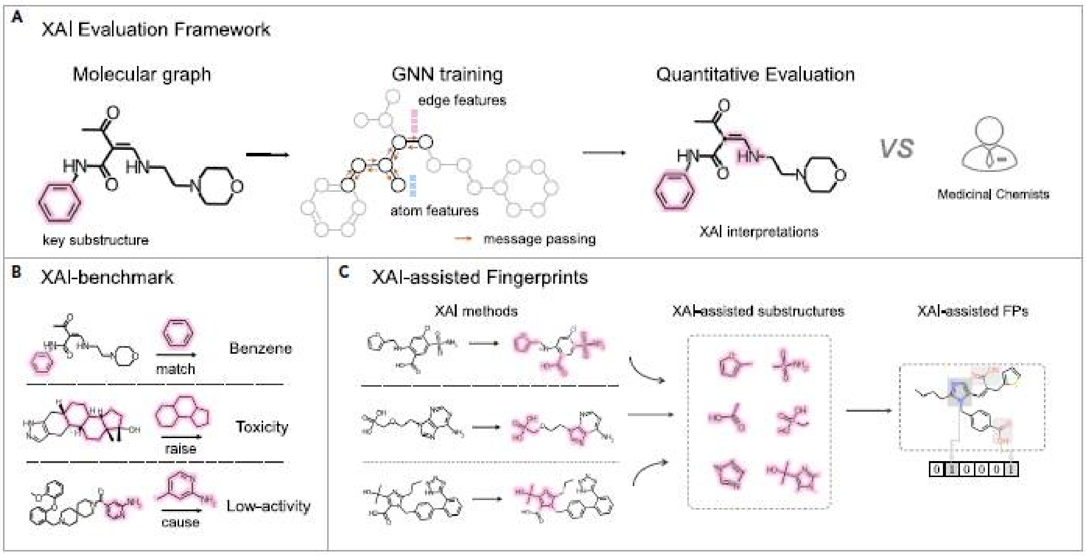
Fit (Figure Image Table) Bases: In continuation of our efforts during last four decades in developing intelligent numerical/ reference/ literal databases, knowledge bits, heuristic rules for hypothesis testing, method selection in Chemical Speciation, Kineto-metrics, Enviro-metrics and Pisci-metrics, here passive data base named Fit. Base incorporating information in xAI-applications is reported.

Med-PaLM2:It is a Medical Large Language model(LLM)of Google. It is the first to perform like an expert (85% accuracy) at the test taken on MedQA data set of US Medical Licencing Examination (USMLE) style questions. This version has 18% improvement from the original Med-PaLM (short for Pathways Language Model). It is also the first AI system reaching 72.5% score on MedMCQA dataset containing Indian AIIMS and NEET medical examination questions.

Keywords: Modelling; xAI; Application disciplines; Medicine (diagnosis, treatment, management), ASD; Mortality model; environment (pollution; water demand), Chemistry Molecular properties; fish; Biochemicals;defence, manufacture/synthesis of chemicals, smart-materials, biochemicals/proteins/genes, Nuclear power plant; Smart cities; Finance service; Health services; Education-Medical;

Chemistry
molecular properties

xAI.	**I(T)O.xAI** **Input operated by Transformer gives Out with xAI methods**	2022-074
Schematic view of the XAI study		

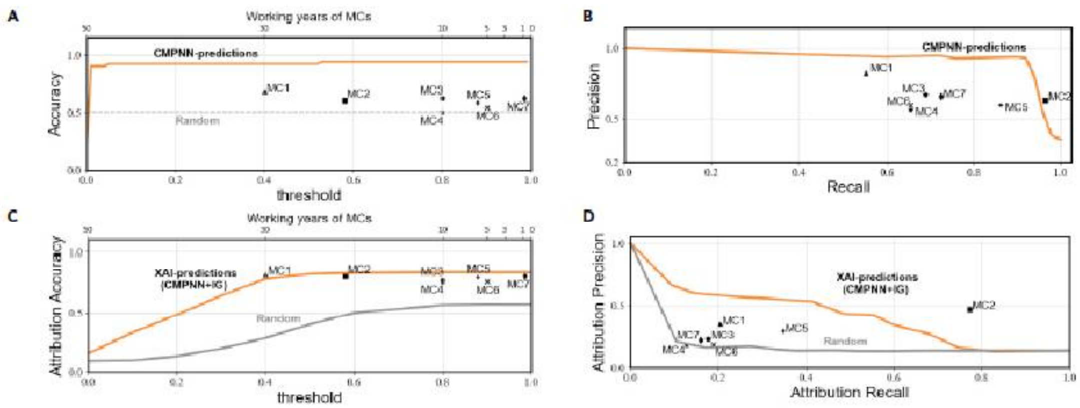


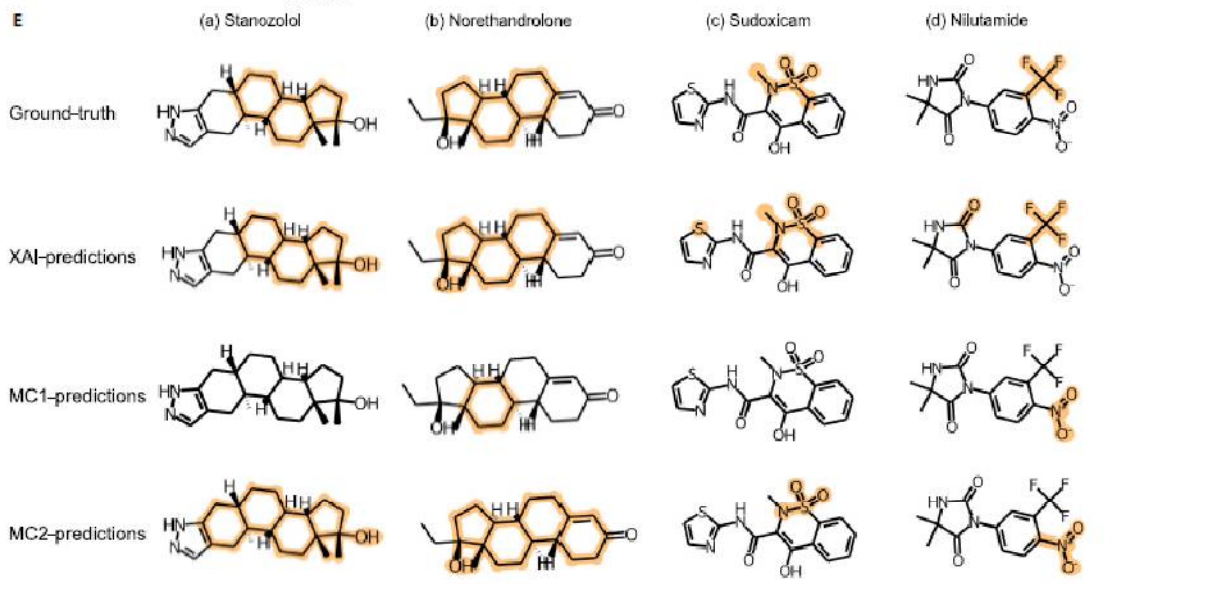
(A) XAI evaluation framework: different prediction models are first trained using state-of-the-art GNN models, which are then interpreted through all XAI methods.
The interpretations are quantitatively assessed and compared with experienced medicinal chemists.

(B) XAI benchmarks including two particular subgraphs (two synthetic benchmarks), the conjunction of multiple substructures, or a local transformation between two molecular graphs (i.e., property cliff).

(C) XAI-assisted fingerprints: the high-frequency key substructures predicted by XAI methods are encoded as fingerprints to input machine-learning models for predicting properties.

Comparison of the XAI (CMPNN + IG) method with medicinal chemists
Accuracy on the hepatotoxicity dataset along the predicted confidence score

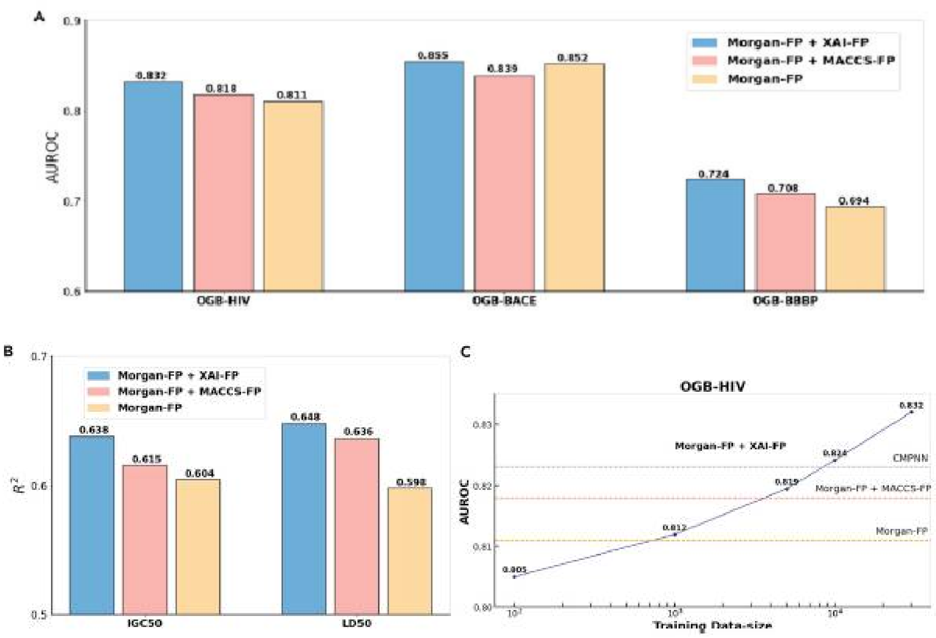




(A) XAI or the working years of medicinal chemists (MCs)
(B) Precision-recall curve
(C) Attribution accuracy for identified substructures
(D) Attribution precision-recall curves
(E) Identified substructures by XAI and MCs

xAI.		2022-074

Performances of Morgan-FP, Morgan-FP + MACCS-FP, and Morgan-FP + XAI-FP by random forest

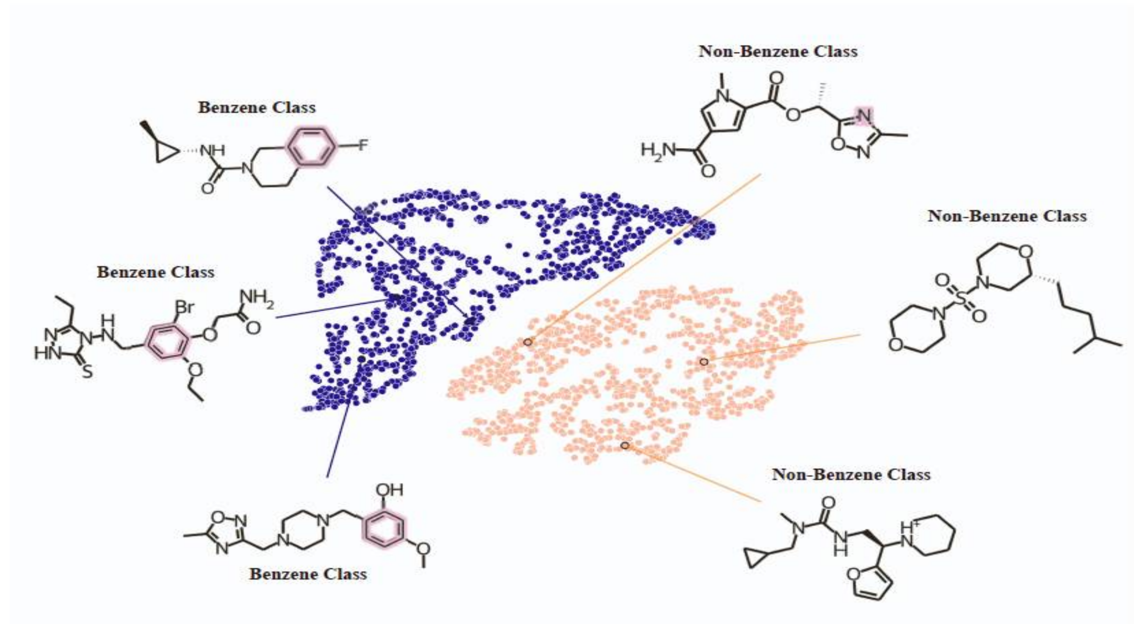


(A) The AUROC of three classification tasks.
(B) The R2 on two regression tasks (IGC50 and LD50).
(C) The AUROC values on the OGB-HIV dataset by Morgan-FP and XAI-assisted fingerprints trained

with different sample sizes

xAI.		2022-074

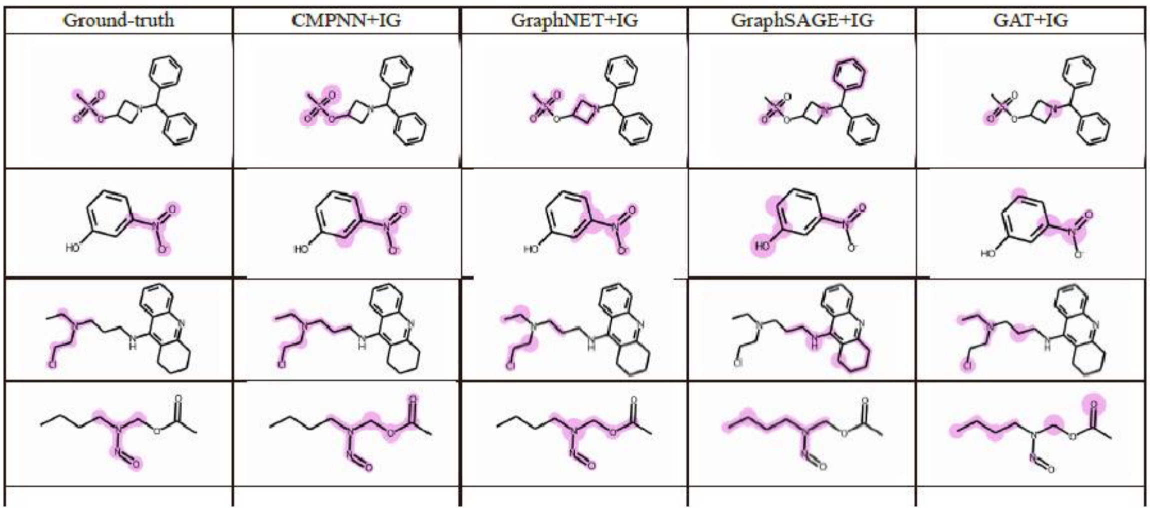
t-SNE Visualization of Benzene dataset



- ✓ Visualization of Benzene dataset. The embeddings learned by CMPNN+IG over the Benzene dataset shown by t-SNE, together with representatively 3 benzene and 3 non-benzenemolecules.

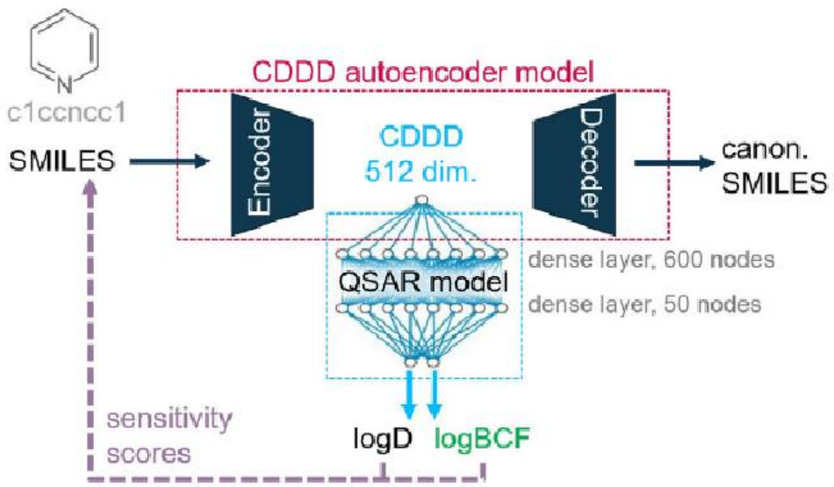
xAI.		2022-074

The predicted substructures by representative combinations of GNN models and XAI Methods

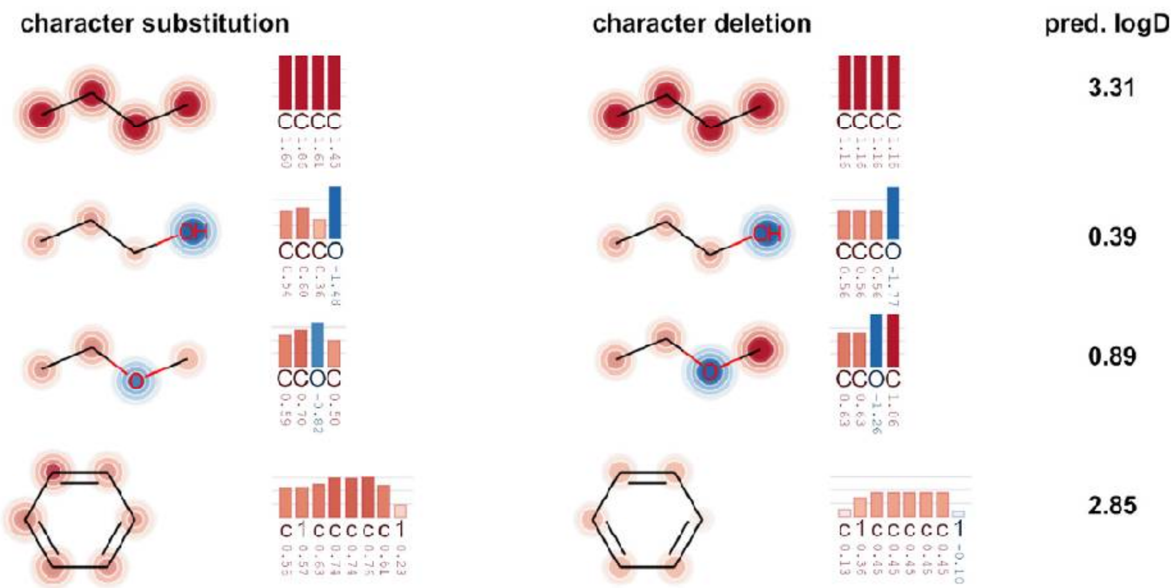


Bi0-Chemical factors in fish

xAI.		2022-005

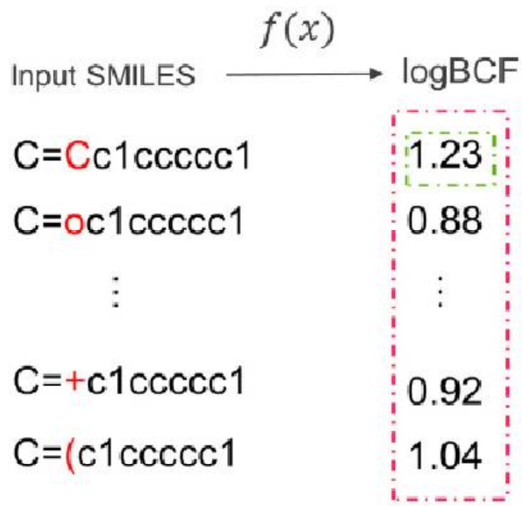


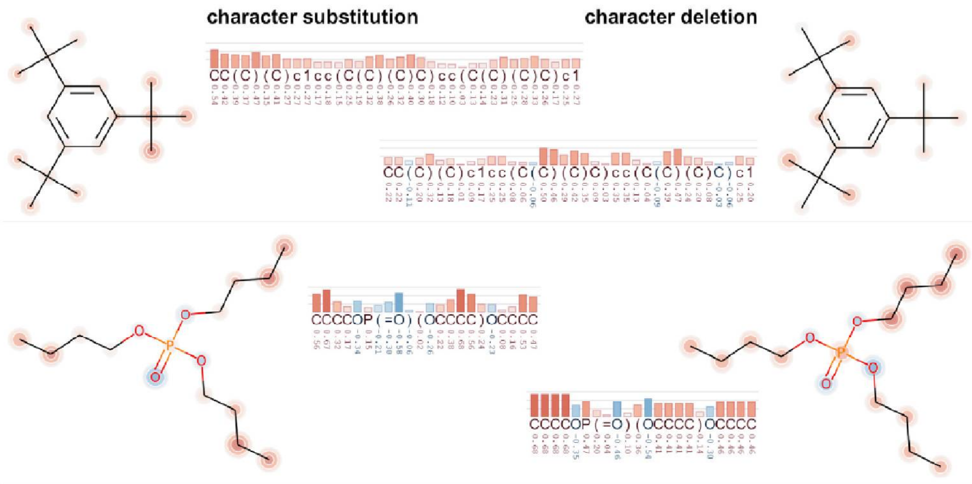
xAI.		2022-005



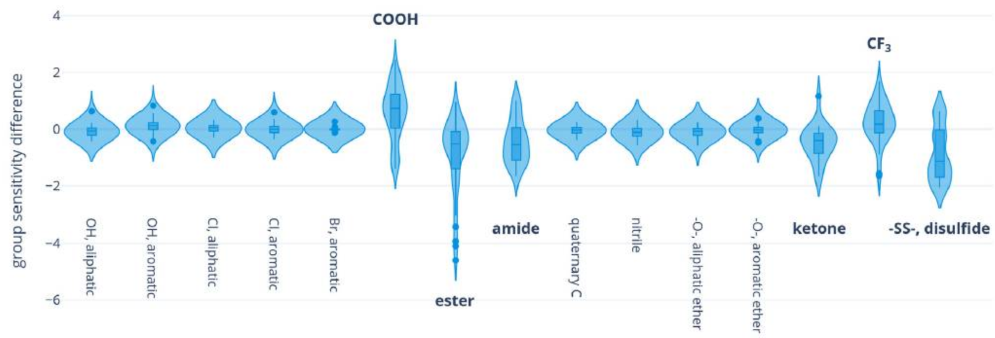
xAI.	Fish	2022-

Modeling bioconcentration factors in fish using xDeepLrn

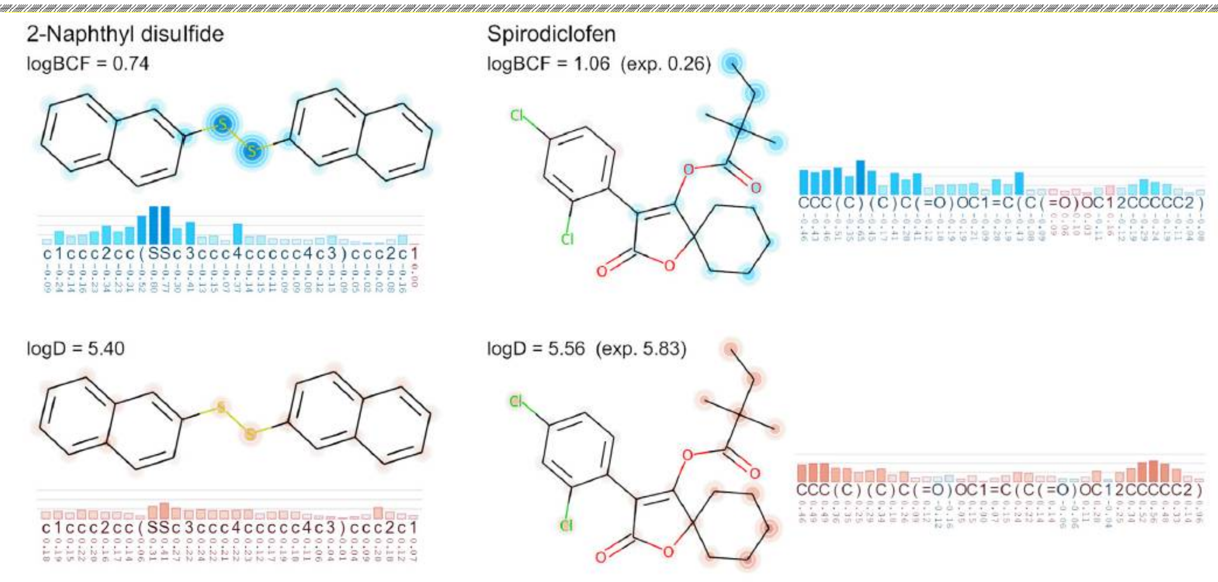




Sensitivity Difference between logBCF and logD for selected functional groups



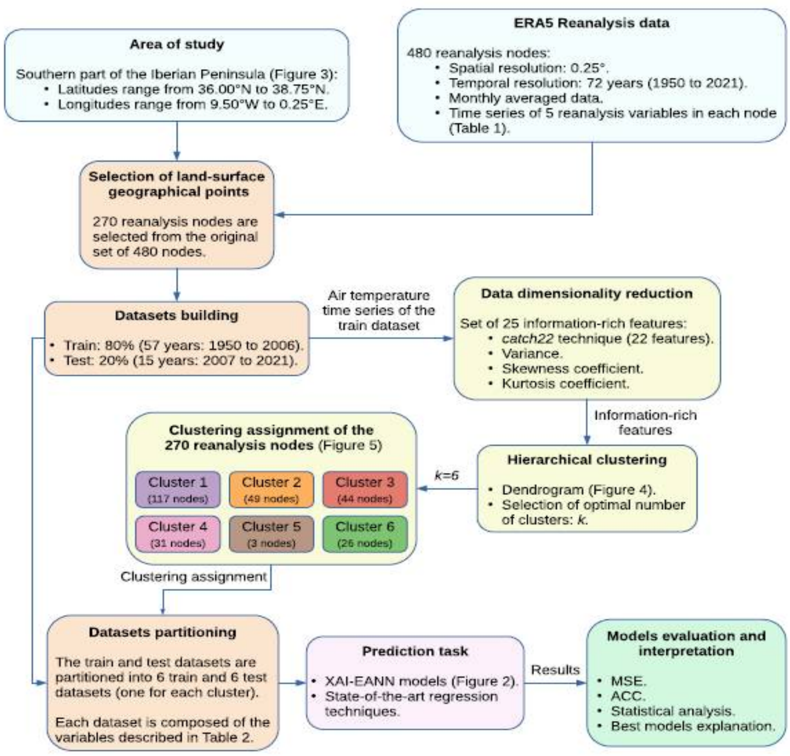
Sensitivity scores for example molecules known to be readily metabolized



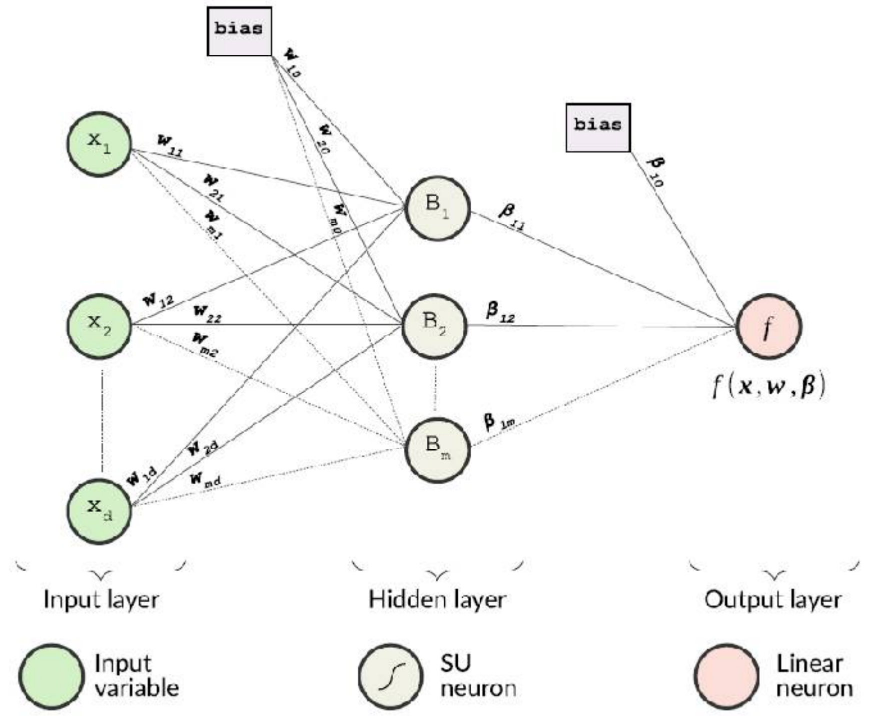
Environment

xAI.	Env, Temp	2023-067

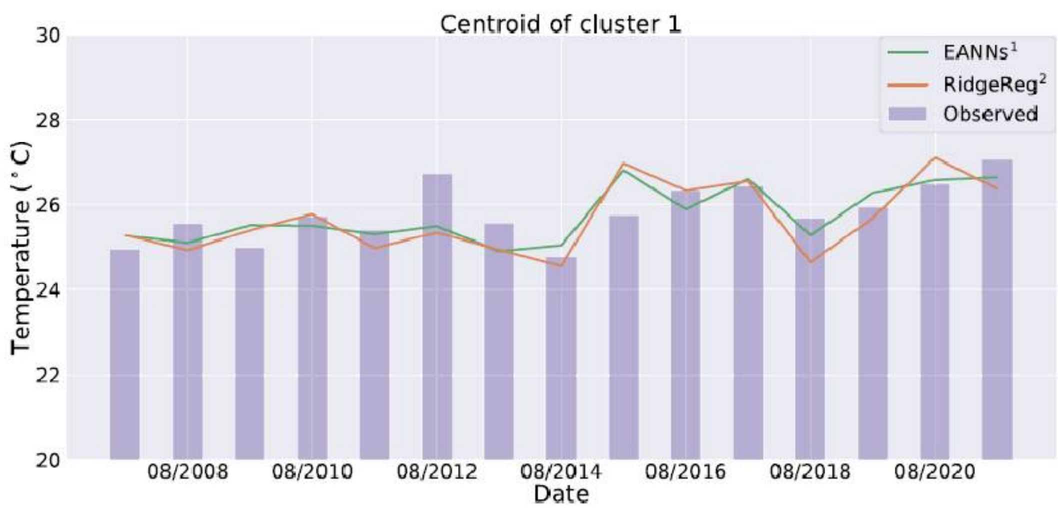
Flowchart of methods for one month ahead prediction of air temperature



Architecture of the ANN regression model



Graphical representation of the observed air temperature values (purple) and the predictions

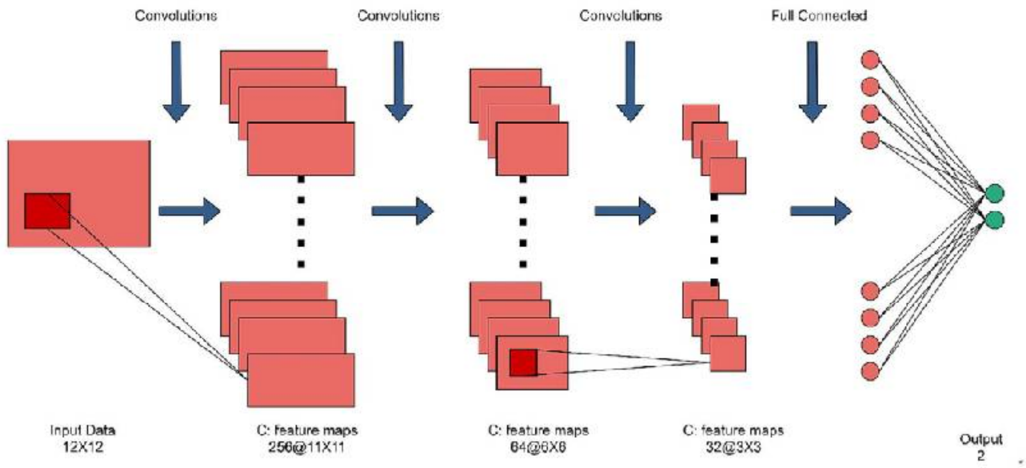


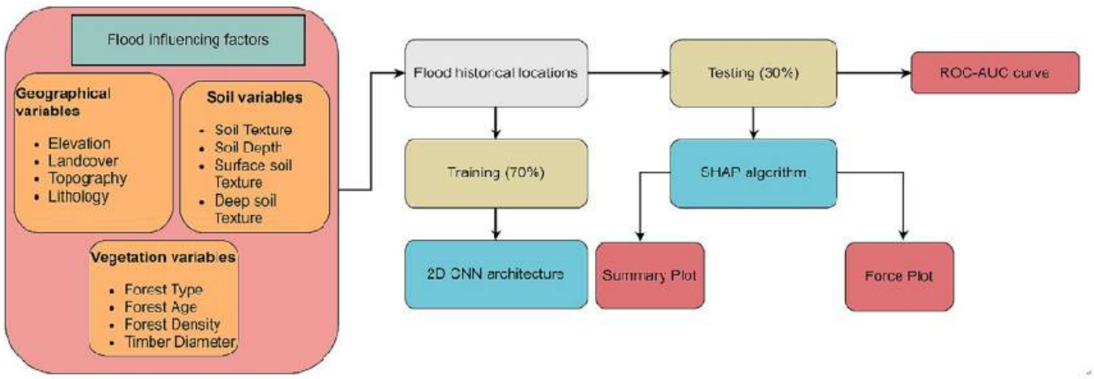
Best XAI-EANN model

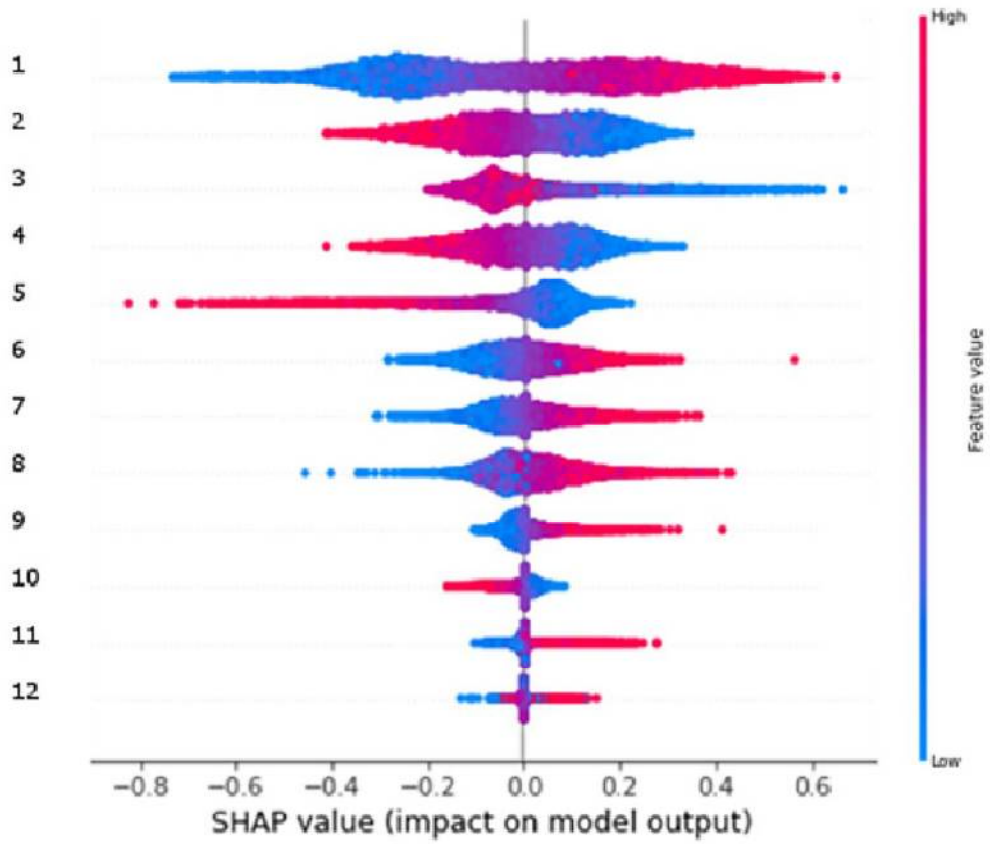
Cluster 1	$\hat{t}_a^* = -0.49 - 8.95B_1 - 6.42B_2 + 5.72B_3 + 4.22B_4$ $B_1 = \sigma(-1.52 + \mathbf{4.79}t^* + \mathbf{3.98}t_{a^*_{-1}} - 1.10v^* - 0.95t^*_{c_6} - 0.04t^*_{c_3})$ $B_2 = \sigma(-2.38 + \mathbf{4.96}t^* - \mathbf{3.27}p^* + 1.24t^*_{c_6} + 0.74v^* + 0.68t^*_{c_5} - 0.66u^* - 0.02t^*_{a_{-1}})$ $B_3 = \sigma(3.76 - \mathbf{5.00}t^*_{c_2} - \mathbf{5.00}t^*_{c_5} - \mathbf{3.87}t^*_{c_4} - \mathbf{3.76}t^* - 1.25u^* + 0.69v^* - 0.32p^*)$ $B_4 = \sigma(5.00 + \mathbf{4.70}t^*_{c_3} - \mathbf{4.44}t^* - \mathbf{3.73}u^* - \mathbf{3.58}t^*_{a_{-1}} - 1.10t^*_{c_4} + 0.89t^*_{c_2} - 0.85p^* + 0.48s^*)$
Cluster 2	$\hat{t}_a^* = 4.04 - 15.14B_1 - 7.52B_2 + 6.77B_3 + 3.89B_4$ $B_1 = \sigma(-1.83 + \mathbf{5.00}t^* - \mathbf{1.86}t^*_{c_3} + \mathbf{1.02}t^*_{a_{-1}} + 0.30t^*_{c_6} - 0.22t^*_{c_4} - 0.20p^*)$ $B_2 = \sigma(-1.03 + \mathbf{1.58}t^*_{c_5} + \mathbf{1.05}t^*_{c_1} + \mathbf{0.96}t^* - 0.28s^*)$ $B_3 = \sigma(3.28 - \mathbf{5.00}t^*_{c_1} - \mathbf{5.00}t^*_{c_3} - \mathbf{3.91}t^*_{c_4} + \mathbf{1.62}t^*_{c_5} - 0.08p^*)$ $B_4 = \sigma(0.93 + \mathbf{4.96}t^*_{c_4} - \mathbf{4.74}u^* + \mathbf{4.11}t^* + \mathbf{3.07}t^*_{c_3} - \mathbf{2.79}t^*_{a_{-1}} - 1.64s^* - 0.75v^*)$

xAI.		2023-043

2D CNN Env



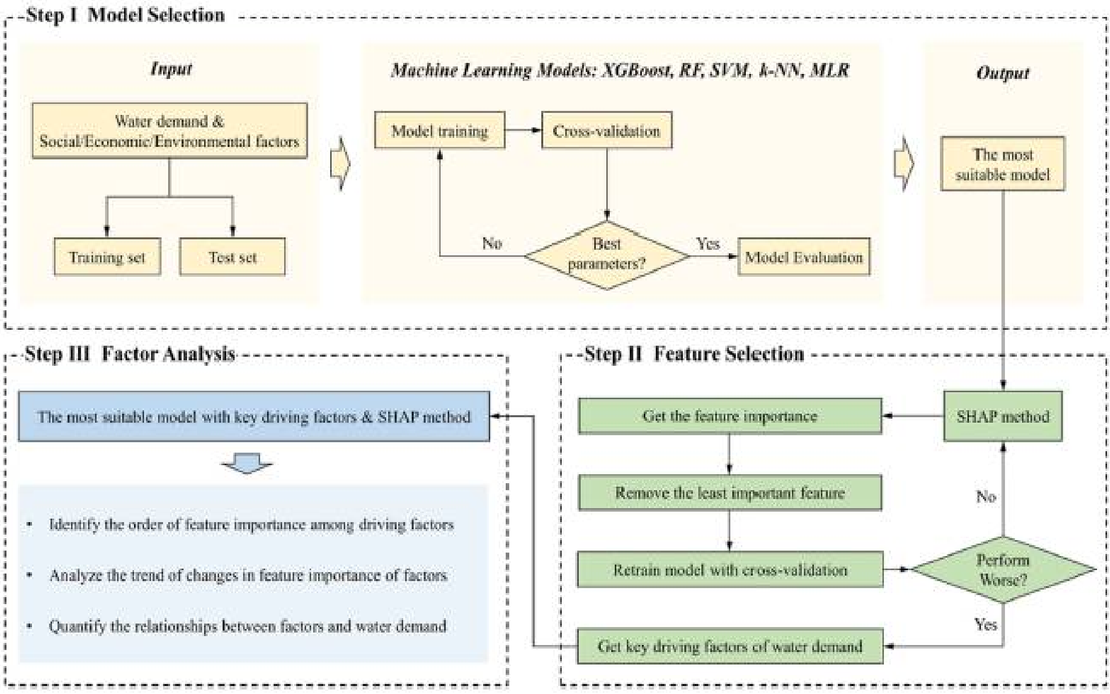




S. No.	Variable	Source
1	Landcover	Ministry of Environment. Korea
2	Elevation	Ministry of Land, Infrastructure and Transport (MOLIT). Korea
3	Soil Depth	National Institute of Agricultural Sciences. Korea
4	Soil Drain	National Institute of Agricultural Sciences. Korea
5	Surface soil texture	National Institute of Agricultural Sciences. Korea
6	Forest Age class	Korea Forest Service
7	Deep soil Texture	National Institute of Agricultural Sciences. Korea
8	Timber diameter	Korea Forest Service
9	Tree Types	Korea Forest Service
10	Forest Density	Korea Forest Service
11	Topography	Ministry of Land, Infrastructure and Transport. Korea
12	Lithology	Korea Institute of Geoscience and Mineral Resources (KIGAM). Korea

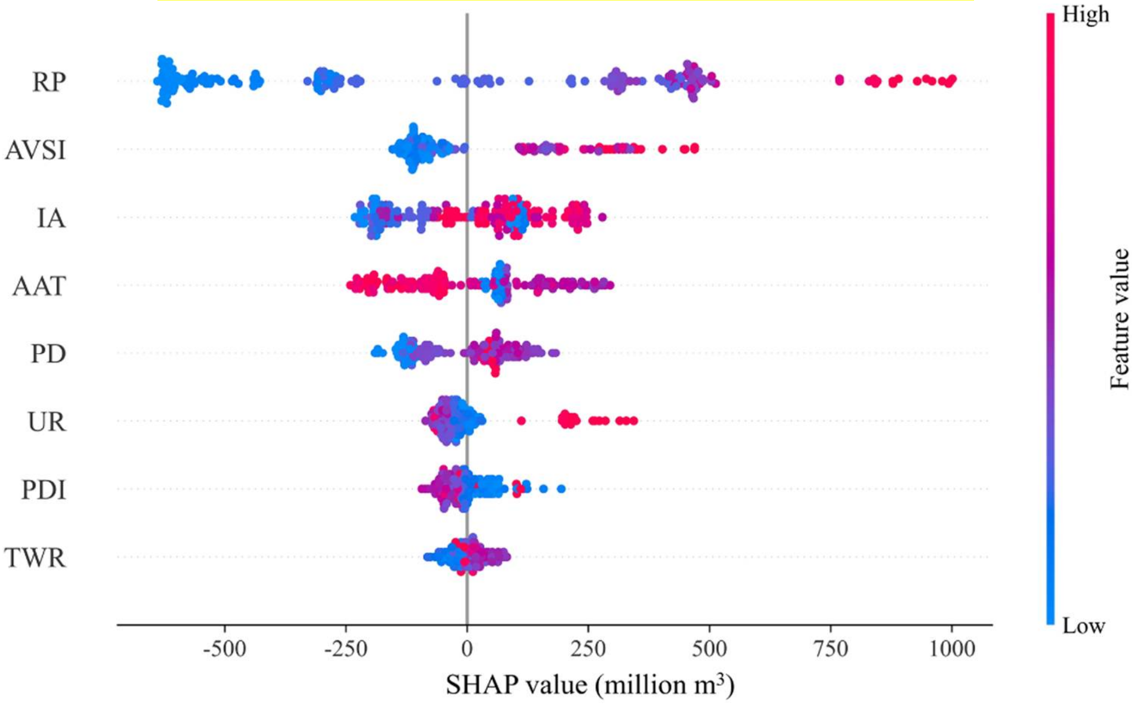
xAI.	Env-water demand	2023-052

xAI-Framework for analysis of factors driving water demand

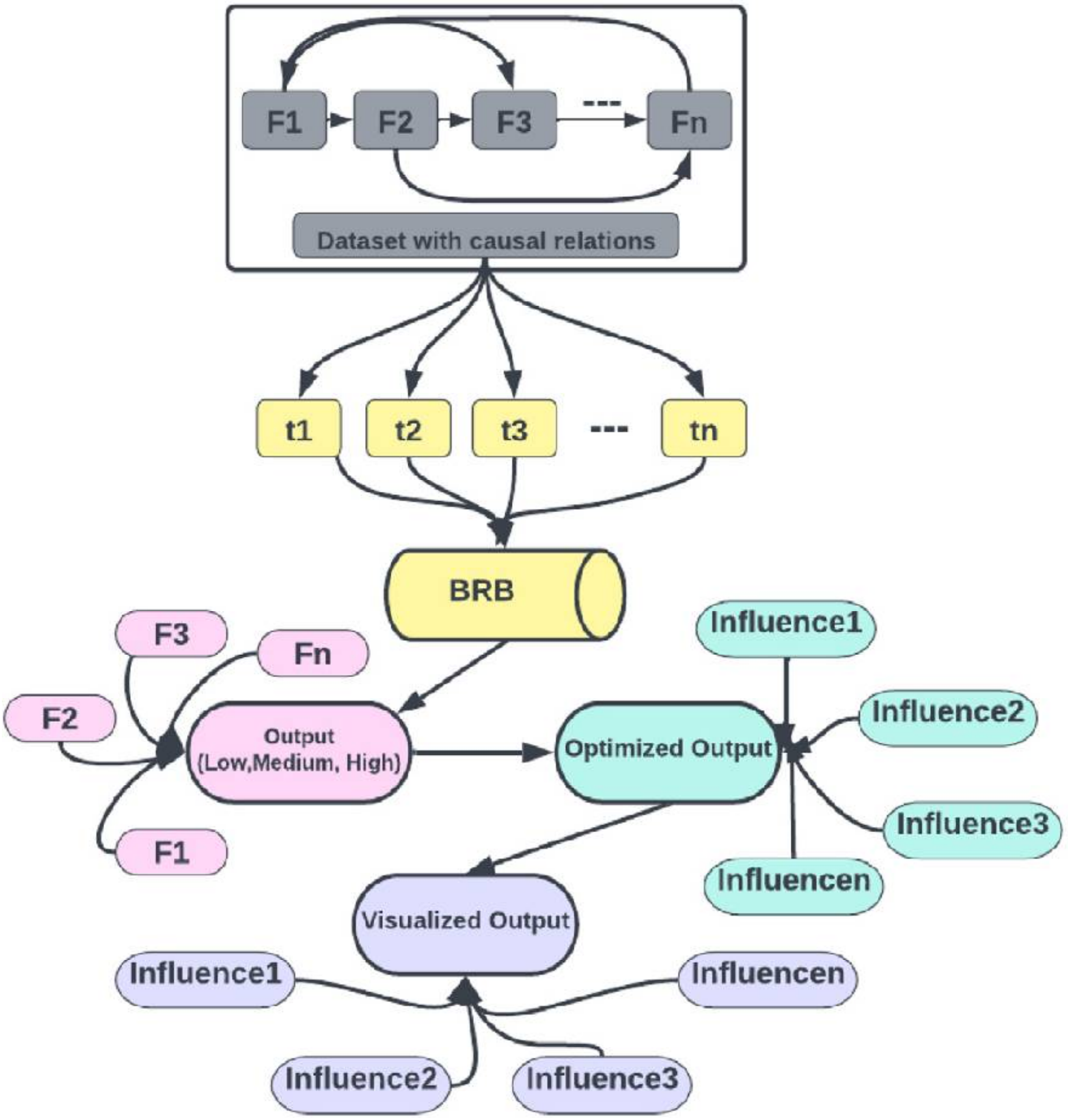


xAI.		2023-052

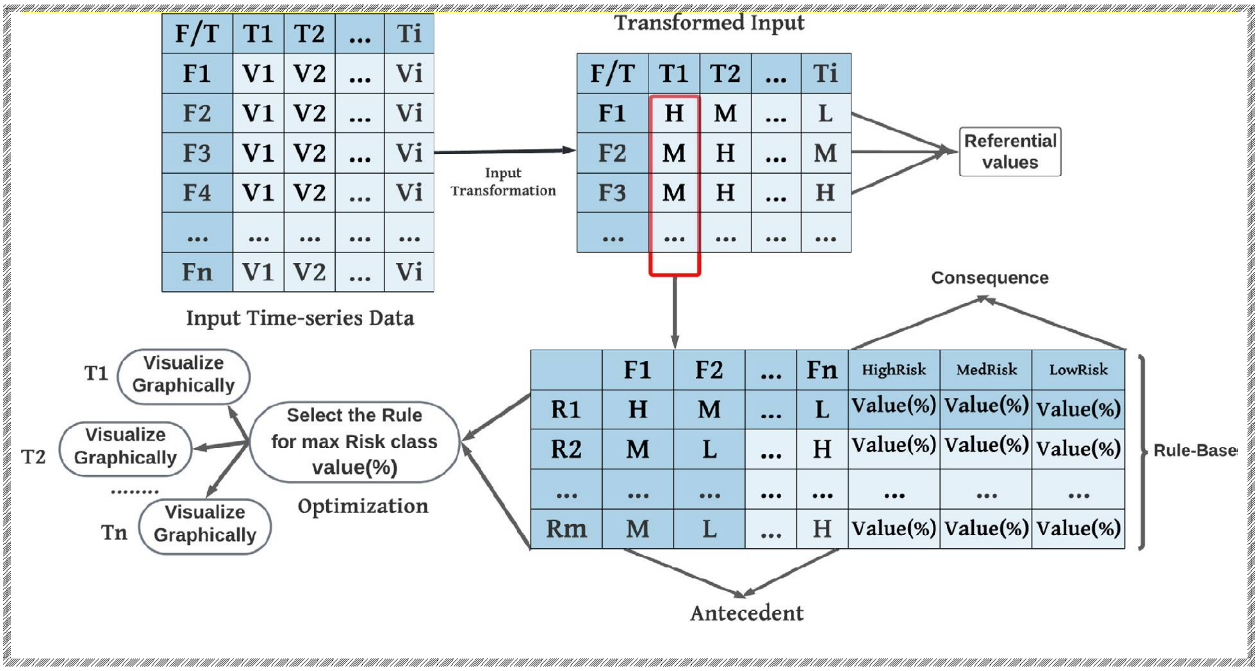
SHapley Additive exPlanations (SHAP) summary plot of input features



Belief-Rule-Based (BRB) architecture



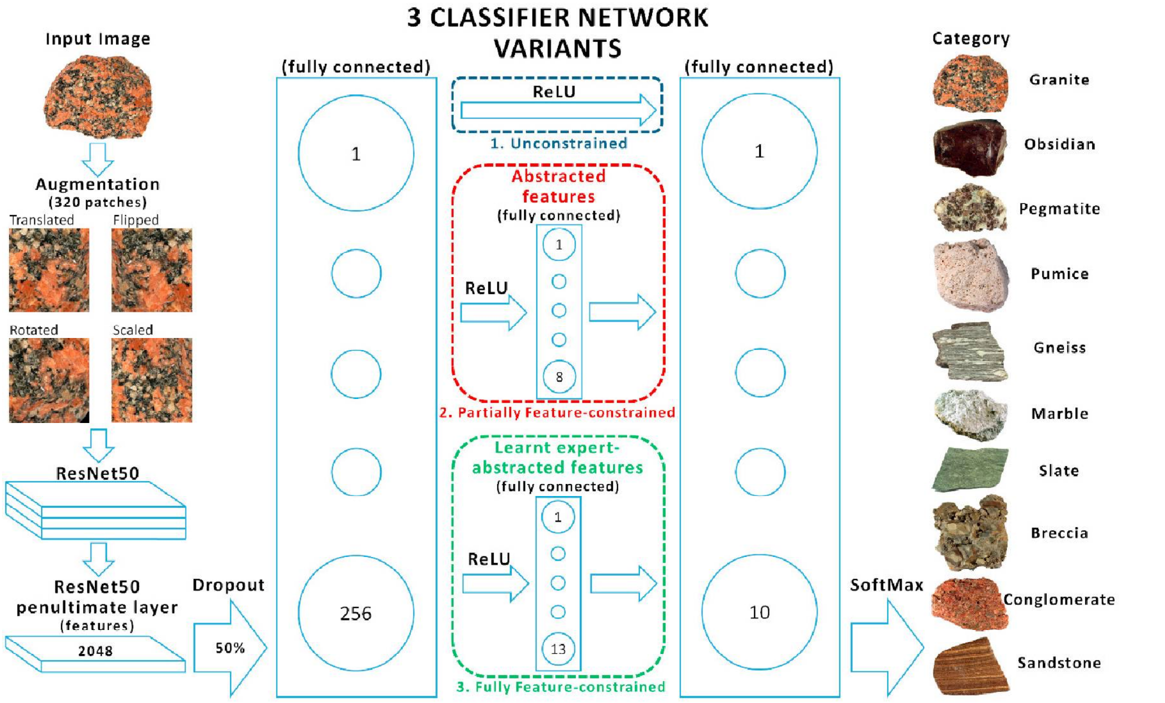
Steps of a BRB Network



xAI.	Rock samples, Env	2022-118

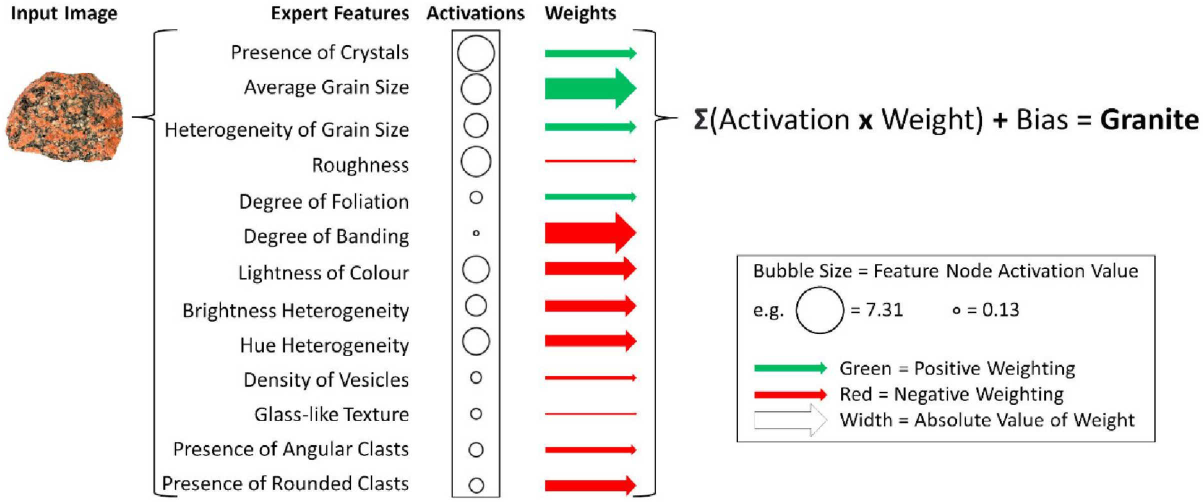
Three classifier network variants

Transfer learning with Renset50



xAI.	Rock samples, Env	2022-118

Categorisation of one instance of the granite images

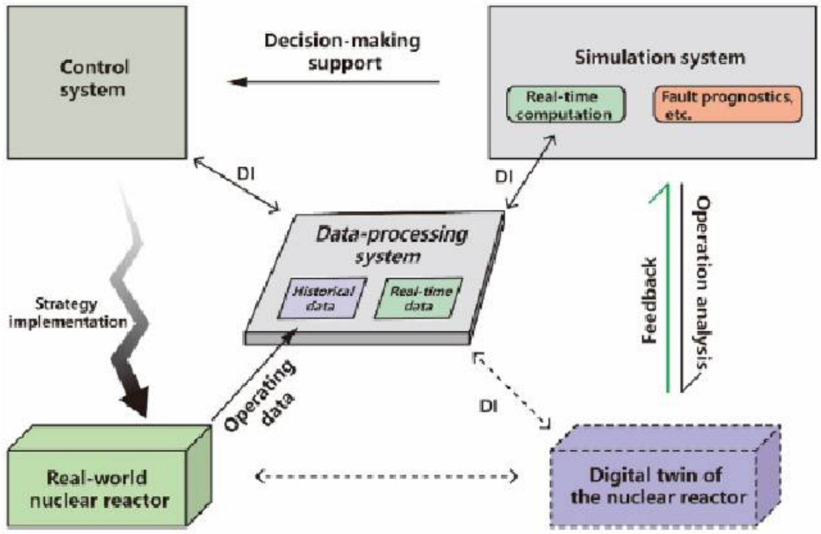


- ✓ Surface area of a bubble: Activations of the Resnet50 transfer-learned expert-abstracted features
- ✓ Width of the arrows: magnitude of weights applied to these feature activations green arrows indicating positive weights and red arrows negative weights

Nuclear power plant

xAI.	Nuclear reactor	2023-039

Interaction operation mechanism of a real-world NPP and its digital twin.

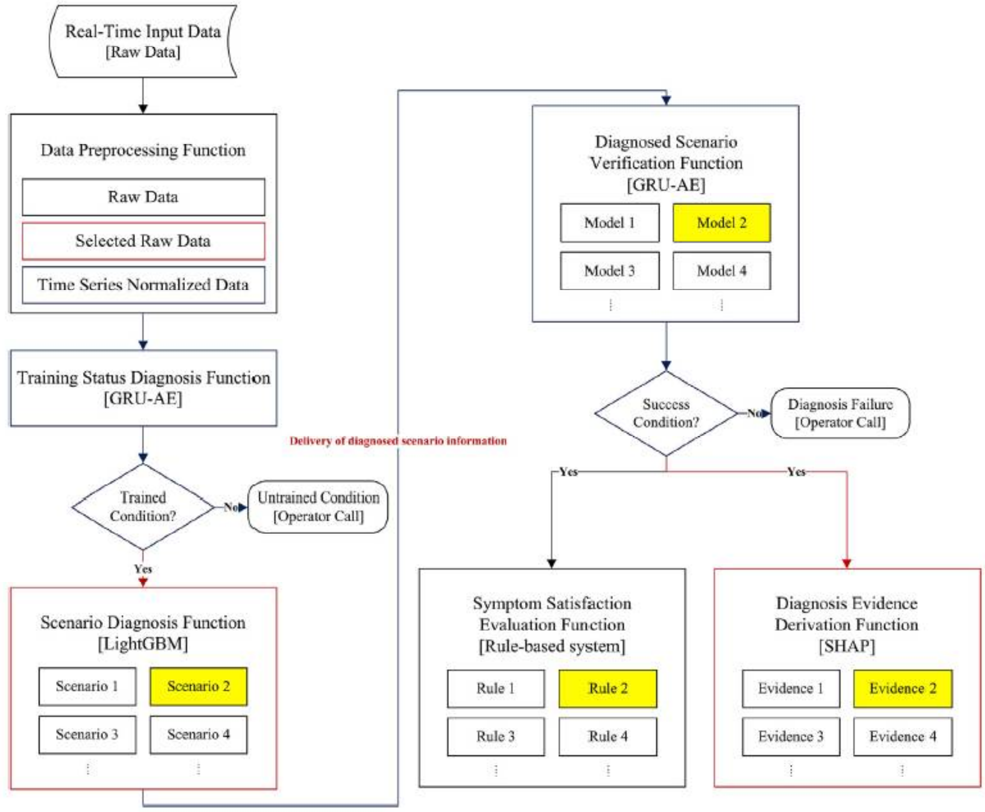


DI: data interaction

xAI.	Nuclear power plant	2022-

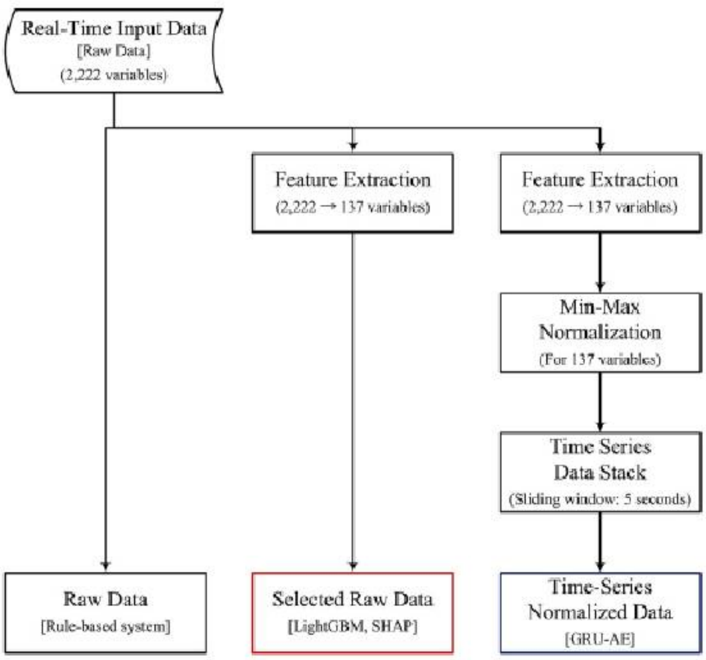
Structure of diagnostic algorithm in the RIDA

(Reliable Intelligent Diagnostic Assistant)



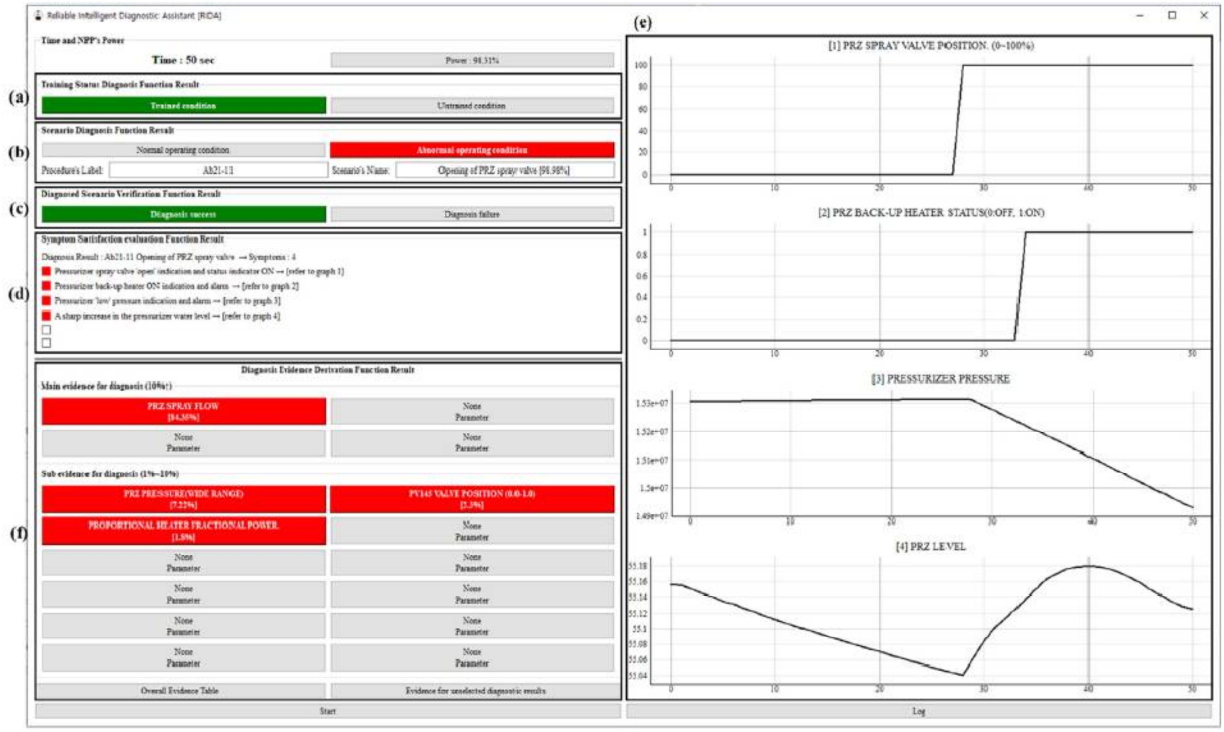
xAI.	Nuclear power plant	2022-

Structure of data preprocessing function



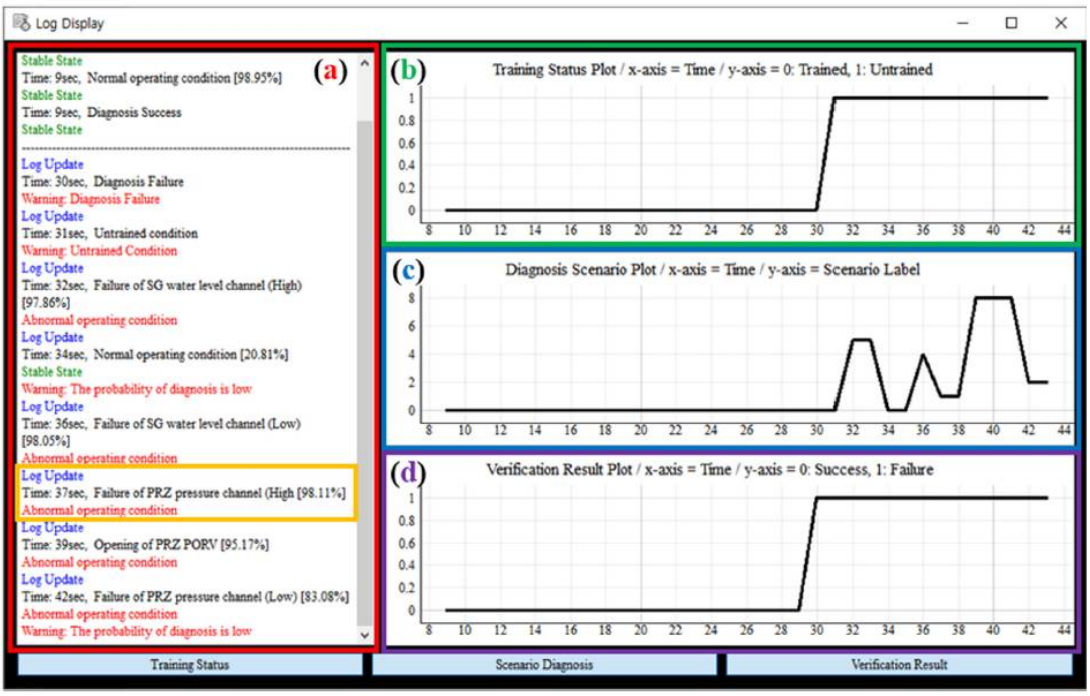
Example of the pop-up window in diagnosis evidence derivation function interface

(a) Graph pop-up window when clicking diagnostic evidence variable (compared with normal operating condition);
(b) Table pop-up window of entire contribution percentage in diagnostic evidence variables;
(c) Pop-up window for providing evidence of undiagnosed scenarios (undiagnosed scenarios are selectable);
(d) Graph pop-up window when clicking undiagnostic evidence variable (compared with selected undiagnostic scenario) Interface result for scenario of PRZ spray valve opening:



(a) The result of training status diagnosis function;
(b) The result of scenario diagnosis function;
(c) The result of diagnosed scenario verification function;
(d) The result of symptom satisfaction evaluation function;
(e) The main symptom variables of symptom satisfaction evaluation function;
(f) The result of diagnosis evidence derivation function

Overview of the log system:

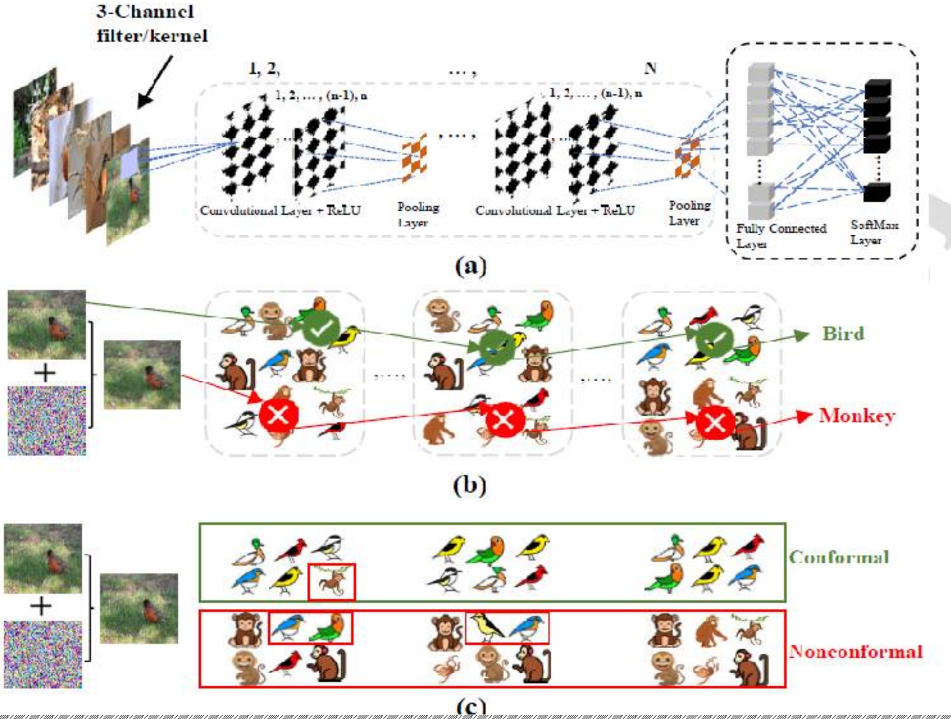


(a) Update contents;
(b) Collected results of the training status diagnosis function;
(c) Collected results of the diagnosis scenario function;
(d) Collected results of the diagnosed scenario verification function

xAI.		2023-030

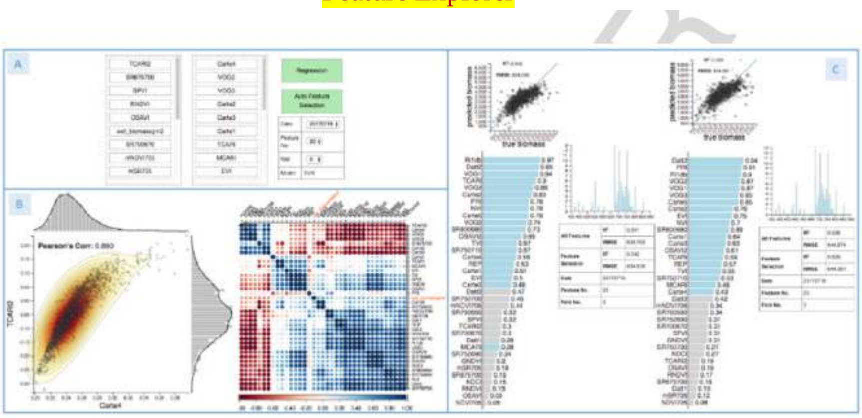
DkNN

- (a) DNN
- (b) Output representations at each layer
- (c) Nearest neighbors at each layer

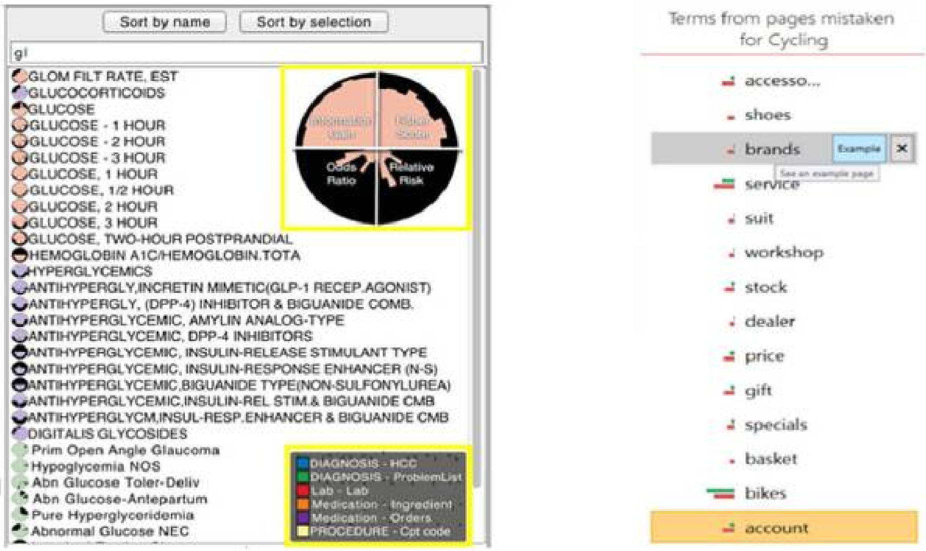


xAI.		2022-

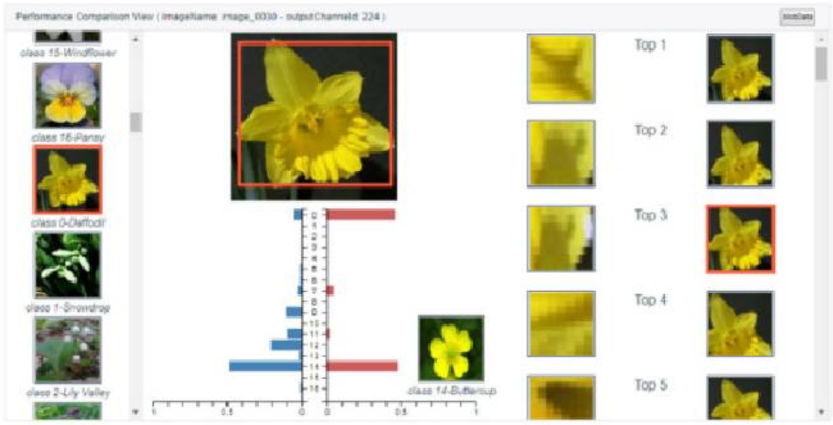
Feature Explorer



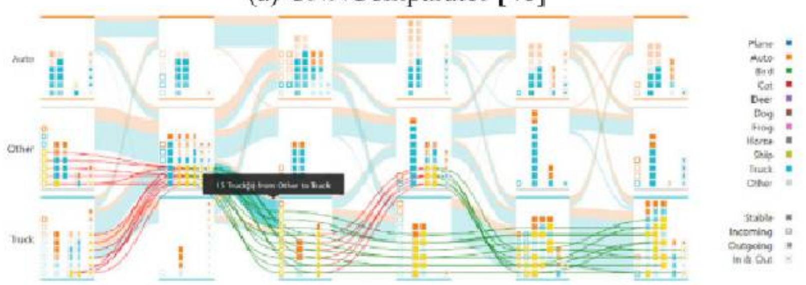
Infuse—Feature insight



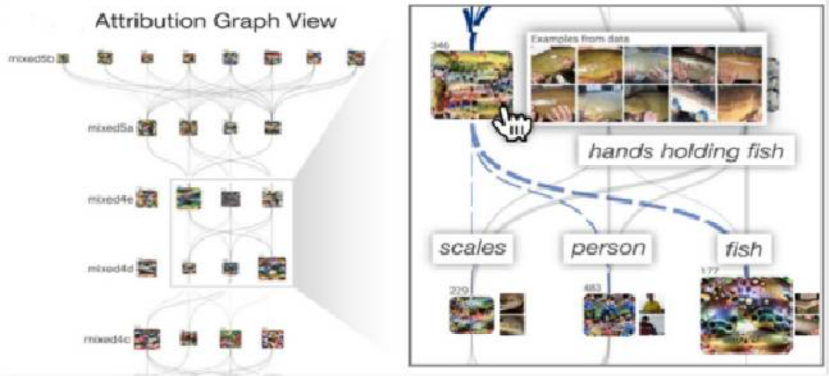
Instance Flow

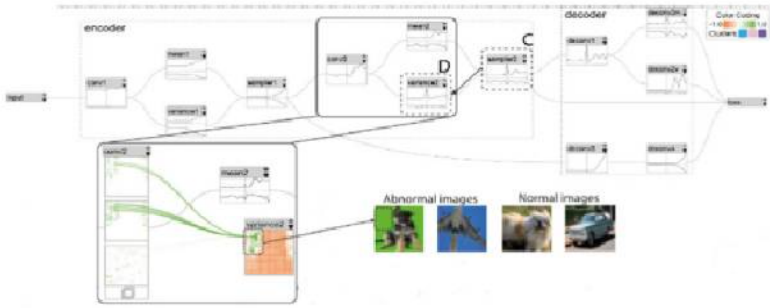


(a) CNNComparator [46]

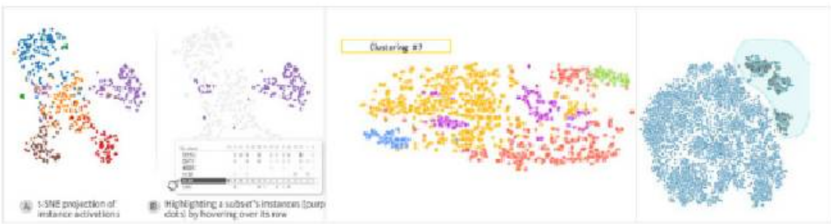


xAI.	2022-
VA tools for understanding architecture of NN models	



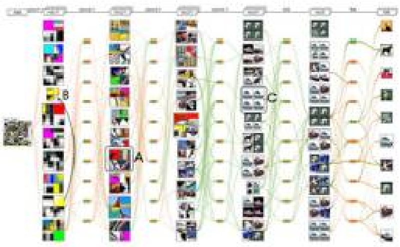


Data representation through scatterplots supported by dimensionality reduction techniques

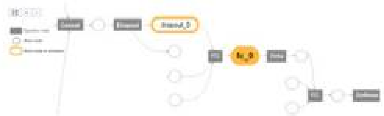


(a) ActiVis [1] (b) ClusterVision [96] (c) RetainVis [59]

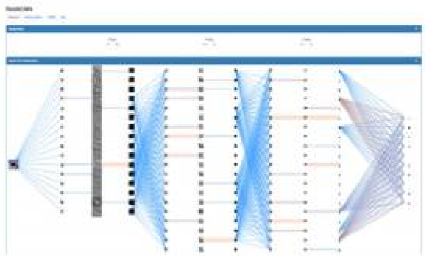
Examples in illustrating NN structure, neurons and hidden layers



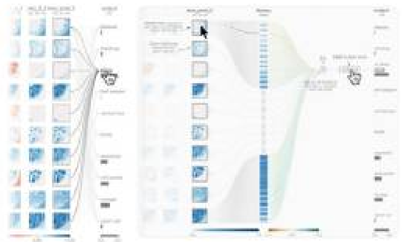
(a) CNNVis [5]



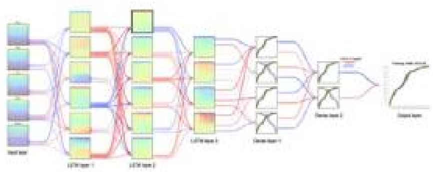
(b) ActiVis [1]



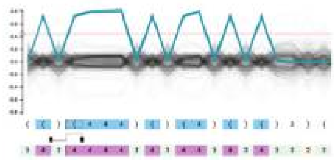
(c) ReCAVNN [9]



(d) CNNExplainer [45]

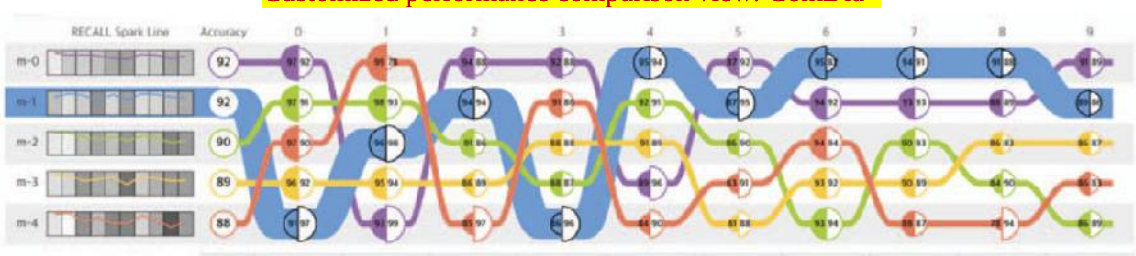


(e) Deepvix [53]



(f) LSTMVis [8]

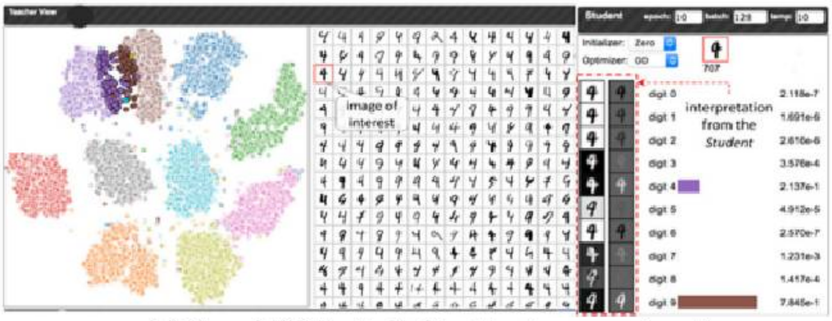
Customized performance comparison view: ComDia+



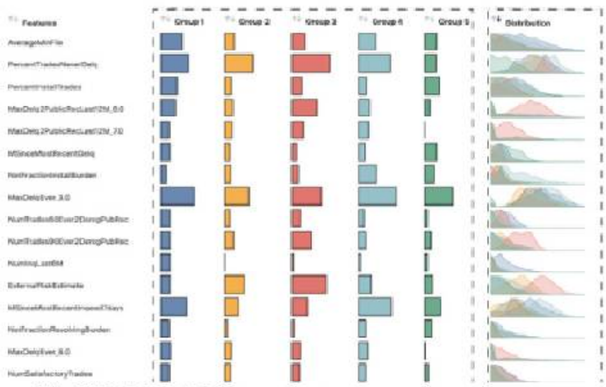
Summary of visual approaches adopted by visual interpretation papers

	Visual approach	Papers
Data representation	Actual data	LSTMVis [8], REMAP [55], CNNExplainer [58], CrossVis [50], DeepCompare [51], Squares [53], TopoAct [66]
	Scatterplots	ActiVis [1], DeepEyes [60], RetainVis [64], Summit [65]
Architecture understanding	DAG	ActiVis [1], CNNVis [5]
	Node-Link	ReVACNN [9], Summit [65]
	Heatmap + PCP	LSTMVis [8], CNNExplainer [58], DeepTracker [61], Deepvix [62],
Performance analysis	Customized visualizations	ActiVis [1], CNNComparator [48], ComDia+ [49], Squares [53], ConfusionWheel [54],
	Traditional visualizations	ReVACNN [9], REMAP [55], CNNSlicer [56], CNNPruner [59], Deepvix [62]

VA tools for both instance and subgroup level explanations by LIME method

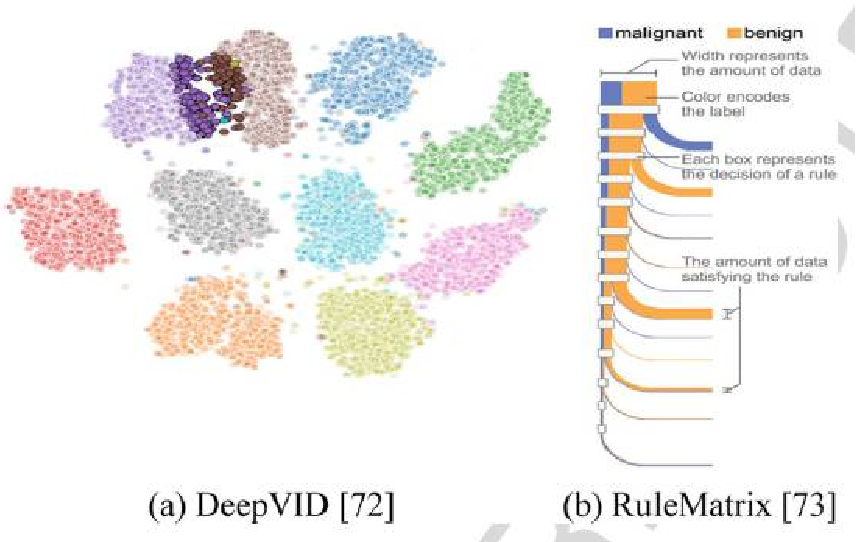


(a) DeepVID for individual instance explanations


(b) SUBPLEX for subgroup explanations

RuleMatrix: If-Then rules visualization in a matrix form

Data representations through scatterplots



(a) DeepVID [72] (b) RuleMatrix [73]

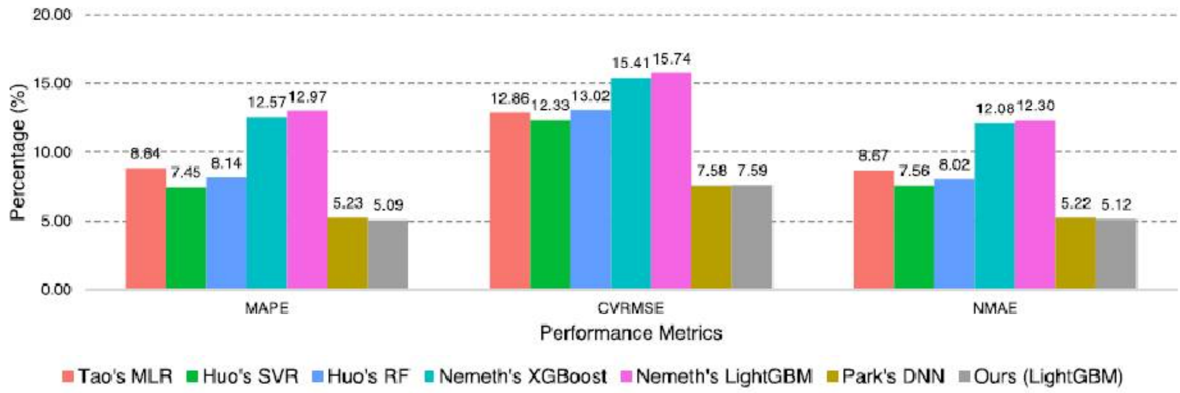
XAI examples

Study	Data representation			Local explanations				Global explanations			
	Actual data	SP	SD	BC	BDP	Heatmaps	O	PCP	H	Matrix	O
ExplainExplore [71]				✔				✔			
SUBPLEX [87]		✔		✔					✔		
MELODY [86]	✔		✔	✔			✔		✔	✔	
explAIner [85]	✔			✔		✔			✔		
RuleMatrix [67]	✔		✔	✔						✔	
DeepVID [72]	✔	✔				✔					
Krause et al. [84]				✔						✔	
iForest [68]	✔	✔		✔			✔				✔
Li et al. [73]		✔		✔							
Botari et al. [75]		✔			✔						
Baptista et al. [79]		✔			✔						
So [83]					✔						
Lamy et al. [80]		✔					✔				
Cho et al. [90]							✔			✔	
Lauritsen et al. [88]		✔					✔		✔		
J. Li et al. [81]		✔				✔					
Kim et al. [78]	✔					✔					

SP: Scatterplot, SD: Sankey Diagram, BC: Bar Chart, BDP: Breakdown Plot, H: Histogram, O: Other, PCP: Parallel Coordinate Plot.

xAI.		2022-061

Comparison of prediction performanceMetrics:
MAPE, CVRMSE, and NMAE

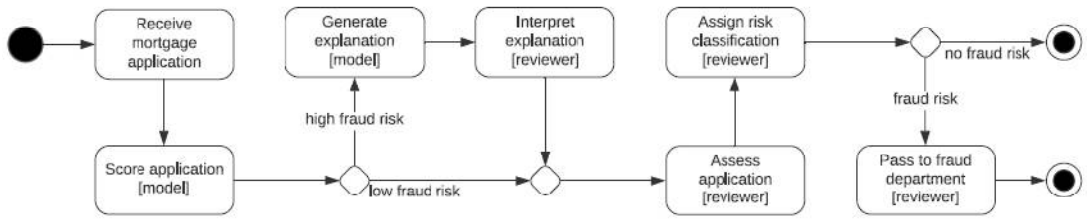


xAI.	Mortgage industry	2022-067

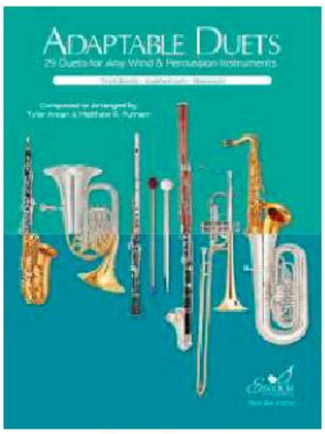
Trust in the model, explanation satisfaction
? Proposed explanation would help the reviewer complete their task

Statement	Median degree of agreement
Trust in the model	
1 - I trust the explanation. I feel like the model is working well.	Agree
2 - I like using the explanations to make decisions.	Agree
3 - I feel like I will make the correct decision only using this explanation.	Neutral
Explanation satisfaction	
4 - The explanations make me understand how the model reaches its judgement.	Agree
5 - I am satisfied with the explanation.	Agree
6 - The explanation is sufficiently detailed.	Agree
7 - The explanation on how the model works seems sufficient.	Neutral
8 - The explanation of the result tells me how accurate the model is.	Agree
Performance	
9 - I reach a decision quicker because the case has an explanation.	Agree
10 - I am able to make a better informed decision because the case has an explanation.	Agree
11 - I find the addition of combinations of risk indicators to the explanation important.	Agree

Activity diagram of MLX1(SHAP and Anchor)



xAI.		2022-



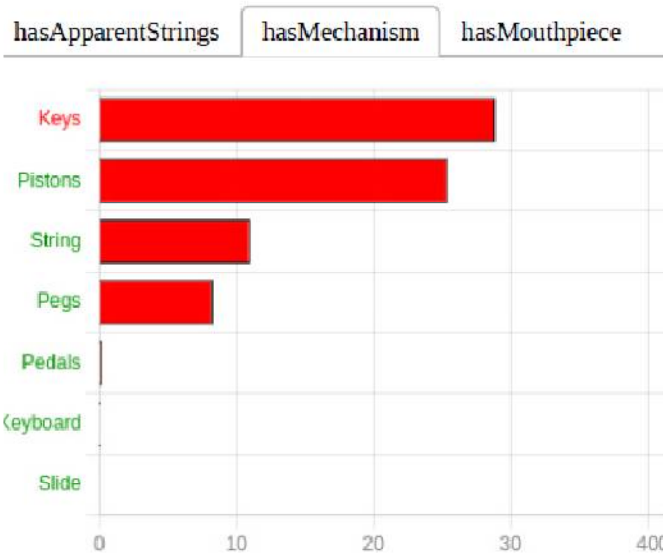
Input image

Bassoon

The model saw a **Bassoon** and it is **not consistent** with the observed properties.

The following properties are inconsistent:

- Not hasMechanism Keys

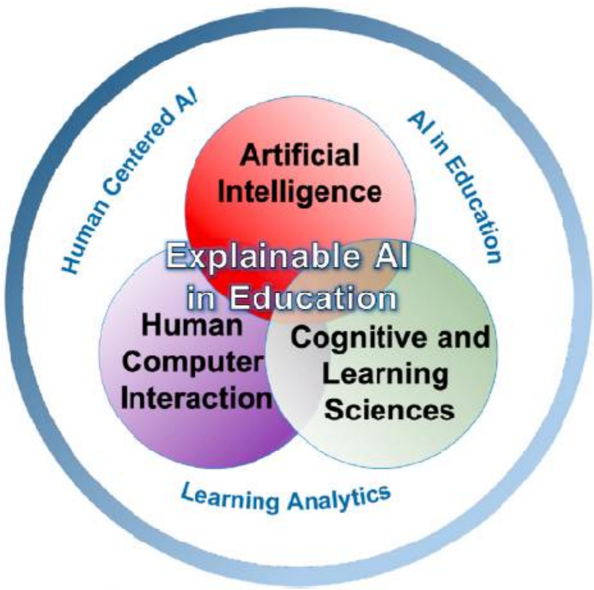


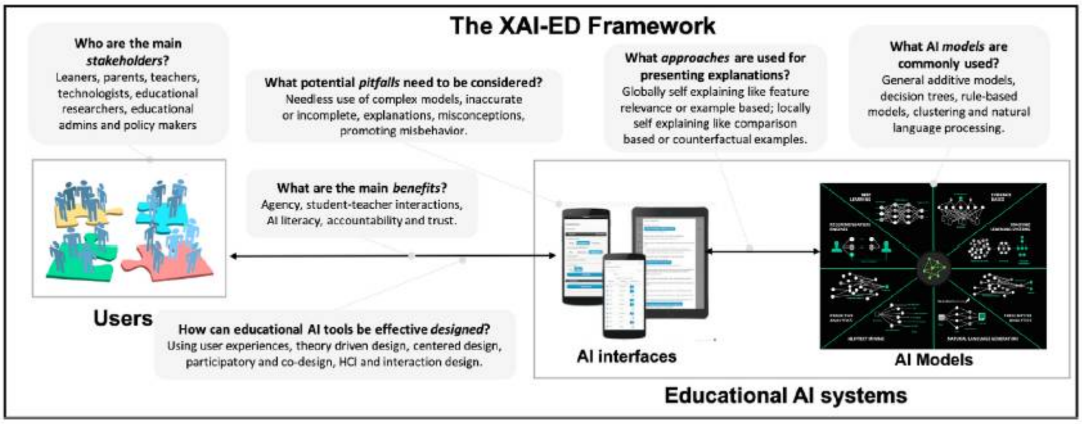
- Green bar: the system has decided the property is present because it has a probability higher than 70 %.
- Red bar: the system has decided the property is absent because it has a probability lower than 30 %.
- Orange bar: the system could not decide about the presence/absence of this property.

xAI. Education

xAI.		2022-172

xAI in Education

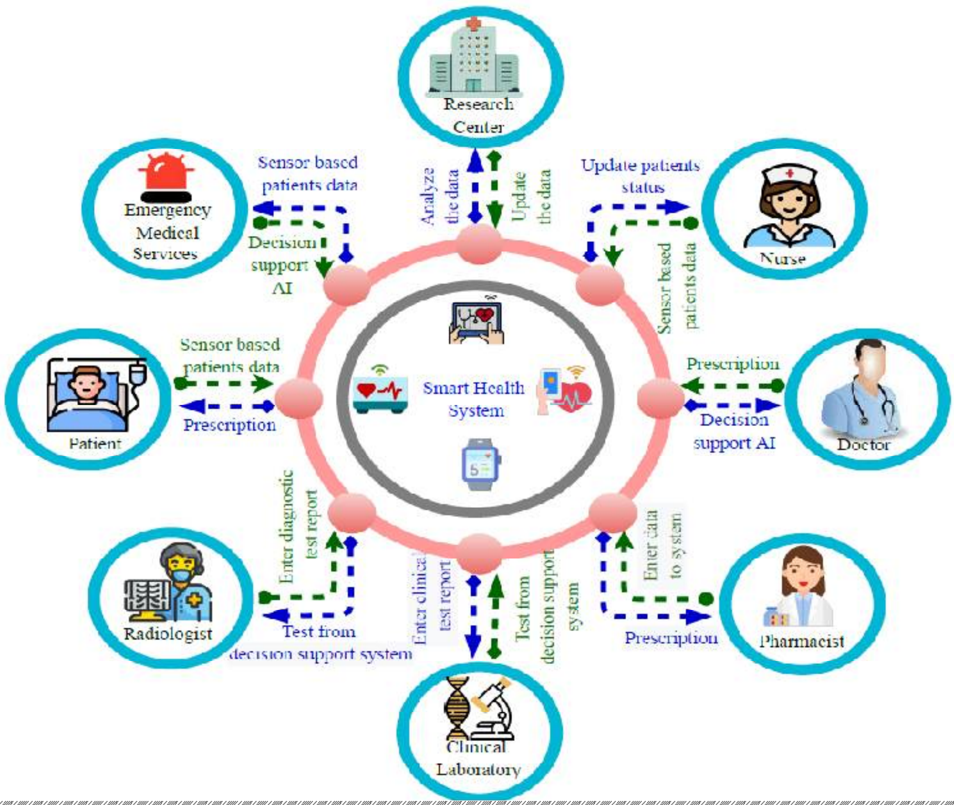




xAI. Medical

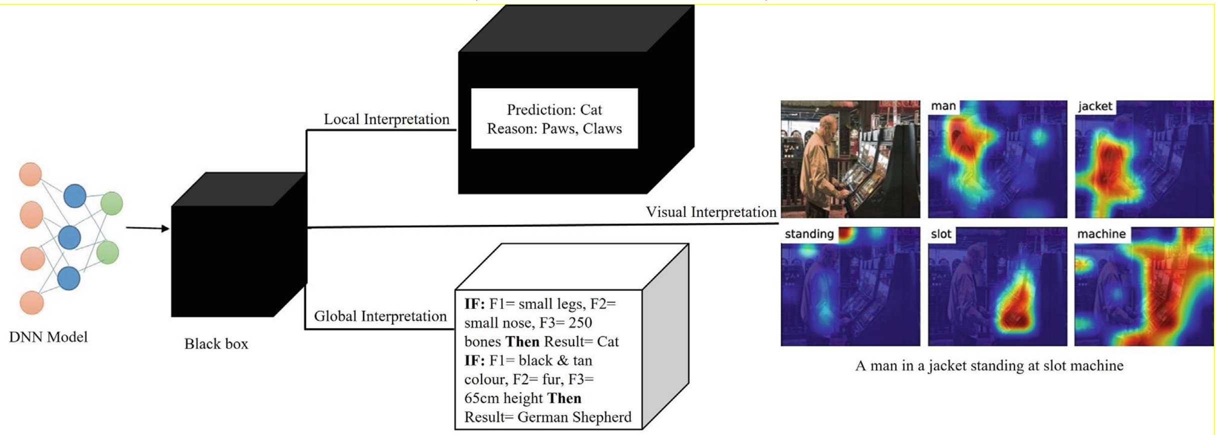
xAI.		2023-141

Smart Health Care System in Smart Cities



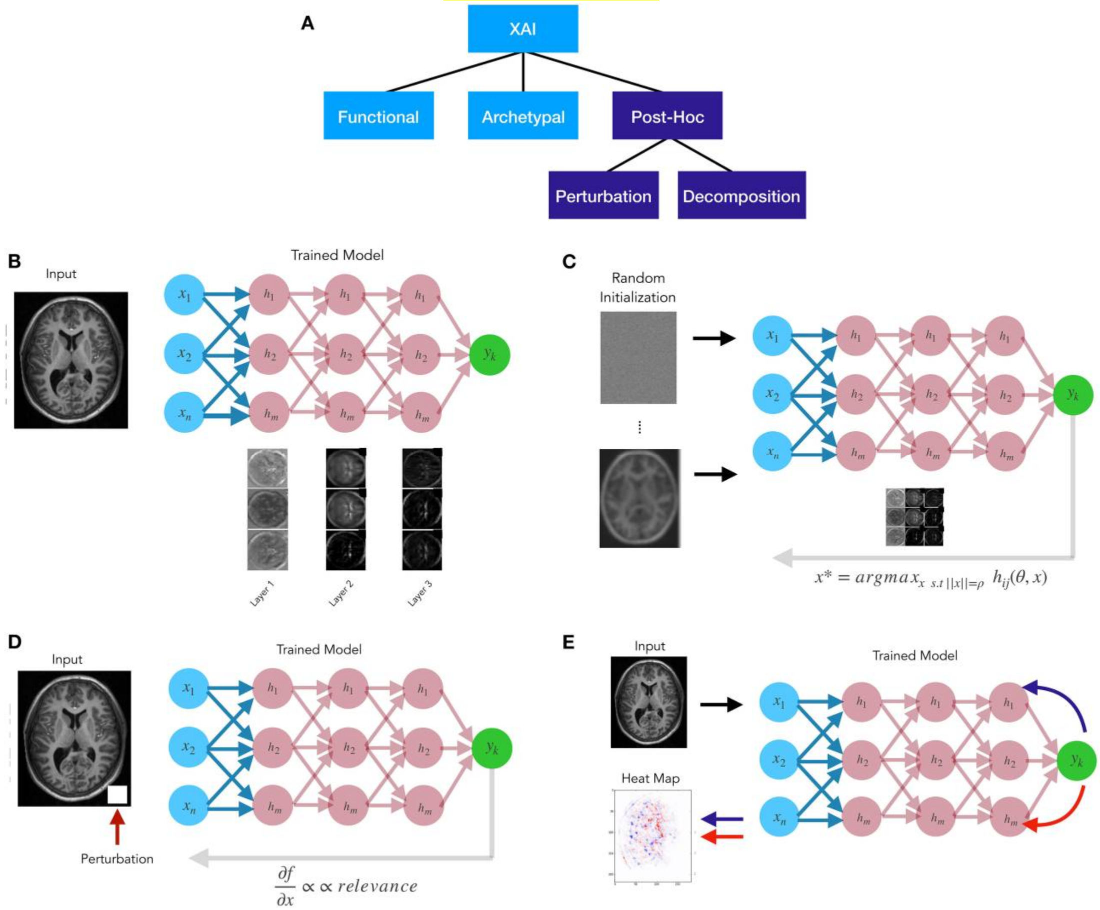
xAI. | 2022-045

Three dimensions (Global, Local and Visual) of XAI methods



xAI. | 2022-103

xAI, Classification



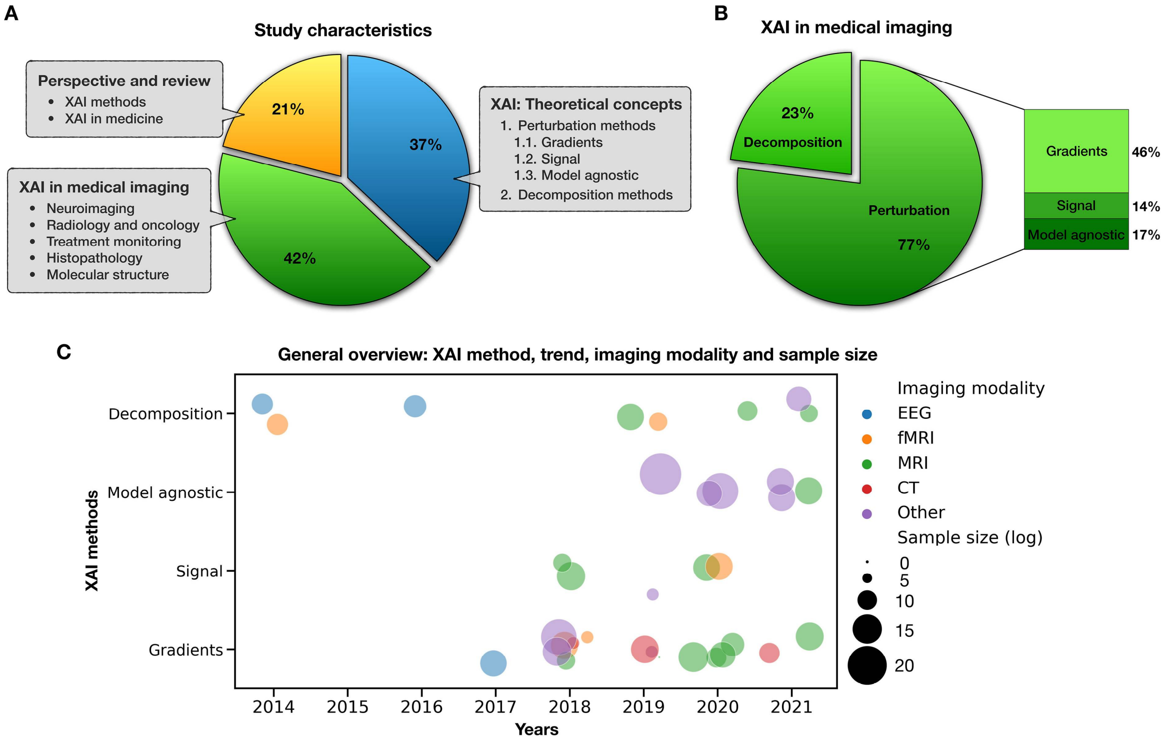
- (A) Explainable AI methods taxonomy.
- (B) Functional approaches attempt to disclose the algorithm’s mechanistic aspects
- (C) Archetypal approaches, like generative methods, seek to uncover input patterns that yield the best model response.
- (D) Post-hoc perturbation relevance approaches generally change the inputs or the model’s

components and then attributing relevance proportionally to the amount of the change in model output

! (E) Post-hoc decomposition relevance approaches are propagation-based techniques explaining an algorithm's decisions by redistributing the function value (i.e., the neural network's output) to the input variables, often in a layer-by-layer fashion.

xAI.	Med	2022-103

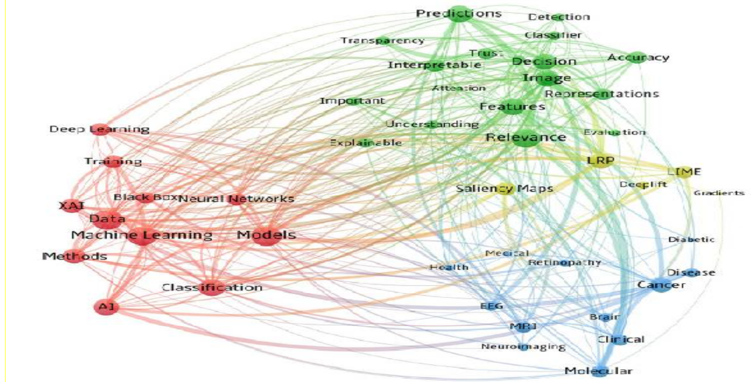
xAI-Med Images



(A) Categorization of included studies
(B) XAI in medical imaging
(C) bubble plot: studies by type of XAI method, imaging modality, sample size, and publication trend in recent years

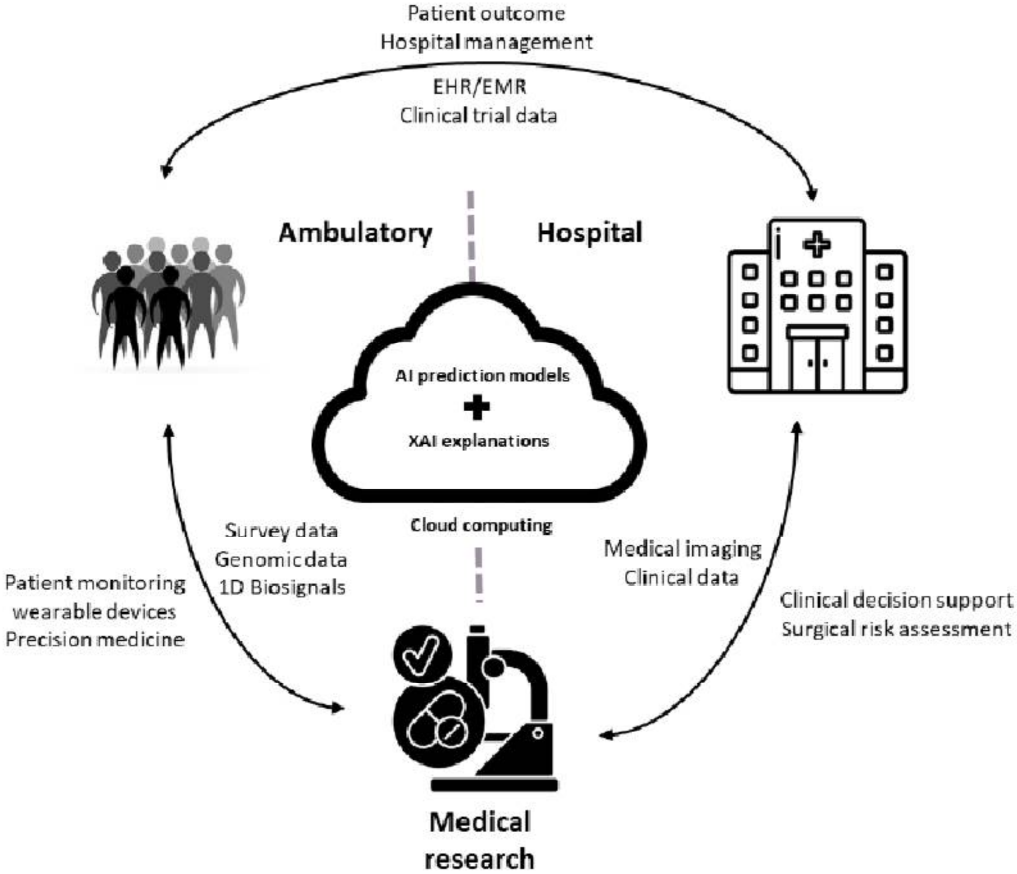
xAI.		2022-103

Co-occurrence network of the commonly used words in reviewed studies

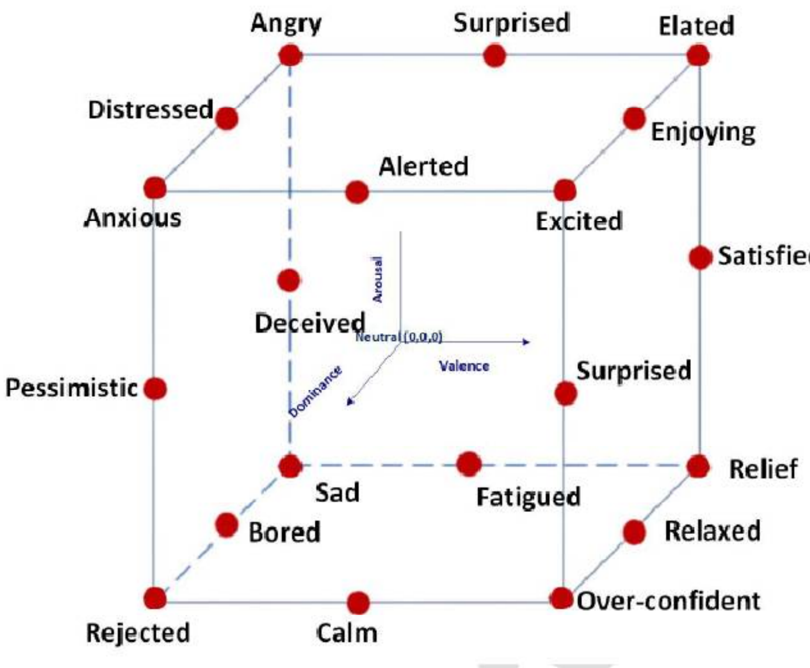


xAI.	Medical	2023-150

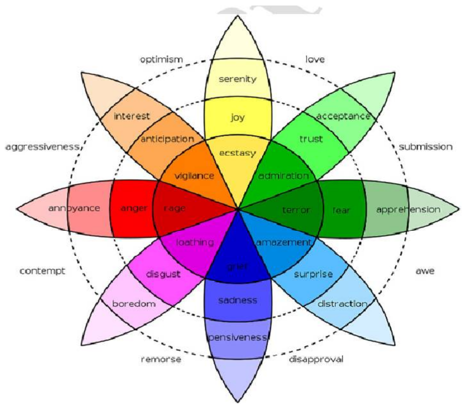
Schematic drawing of cloud system for ambulatory and hospital settings



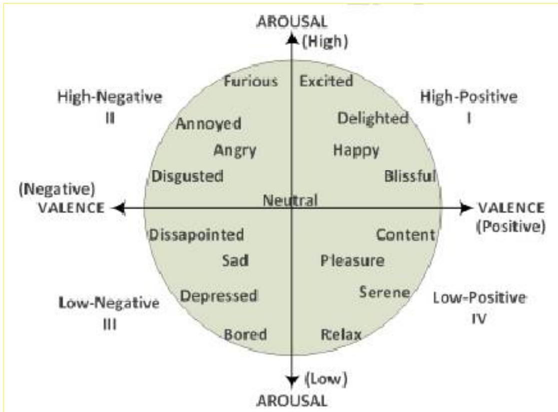
xAI.	Medical, Emotion	2023-161



2D VA emotion model

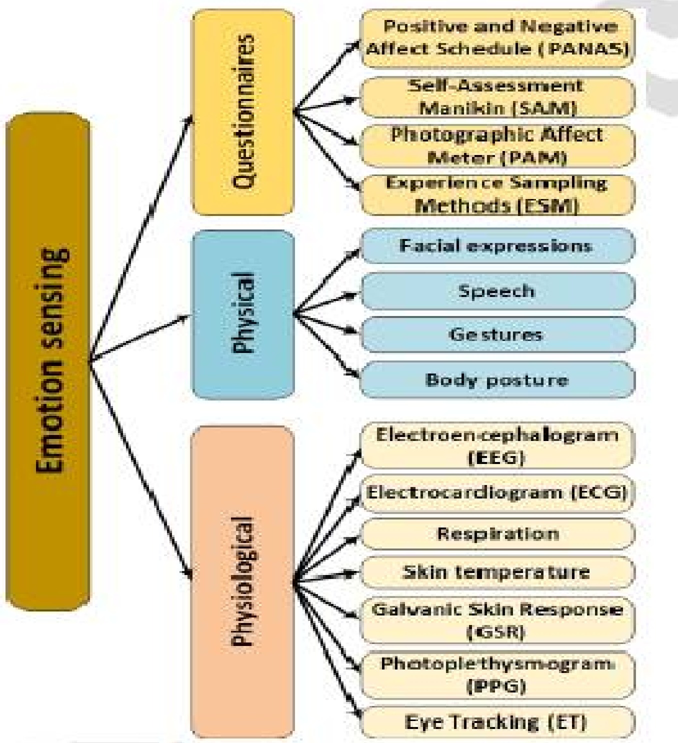


Plutchik's wheel of emotions



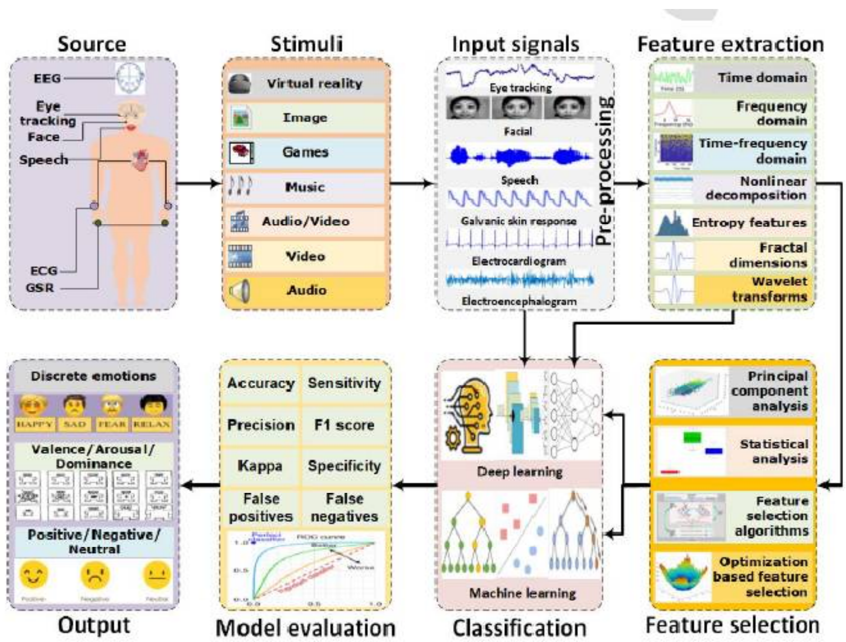
2D VA emotion model

xAI.	Medical, Emotion	2023-161



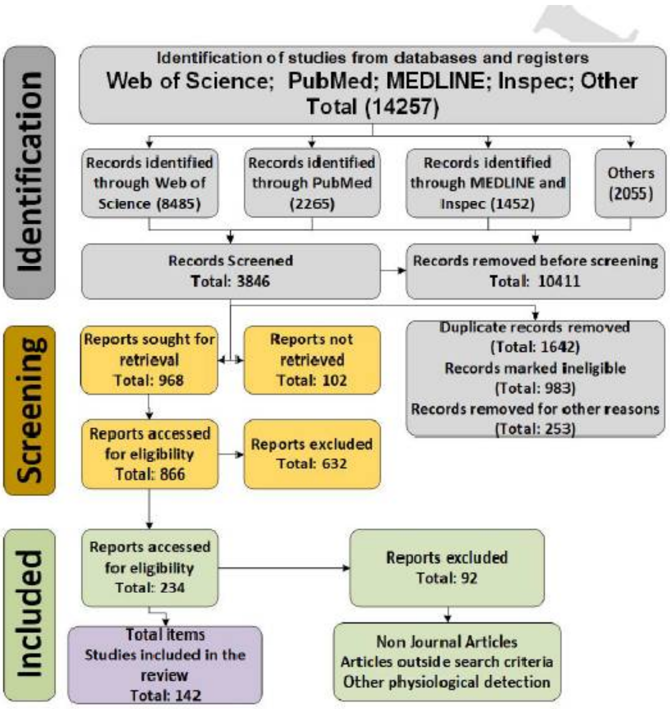
Branching representation and classification of emotion sensing techniques

xAI.	Medical, Emotion	2023-161



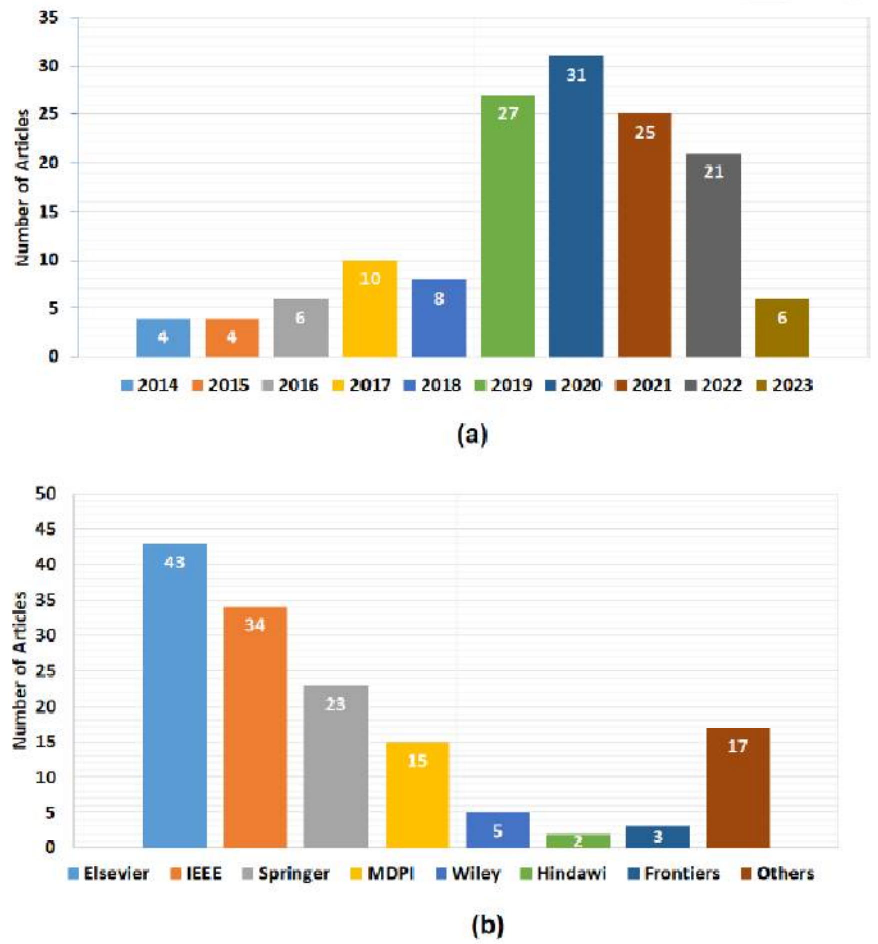
Work flow in an automated emotion recognition system

xAI.	Medical, Emotion	2023-161



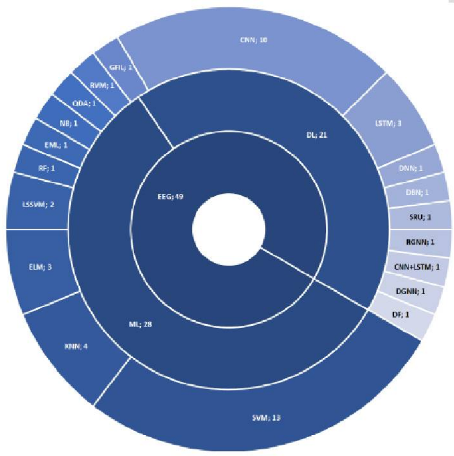
PRISMA guidelines followed during the selection of the articles

xAI.	Medical, Emotion	2023-161

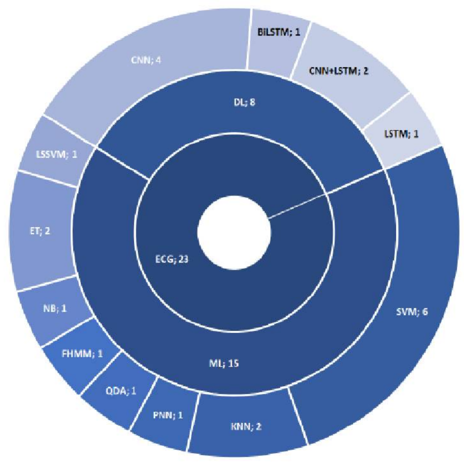


✓ (a) Publisher-based distribution
✓ (b) Year-wise distribution

xAI.	Medical, Emotion	2023-161

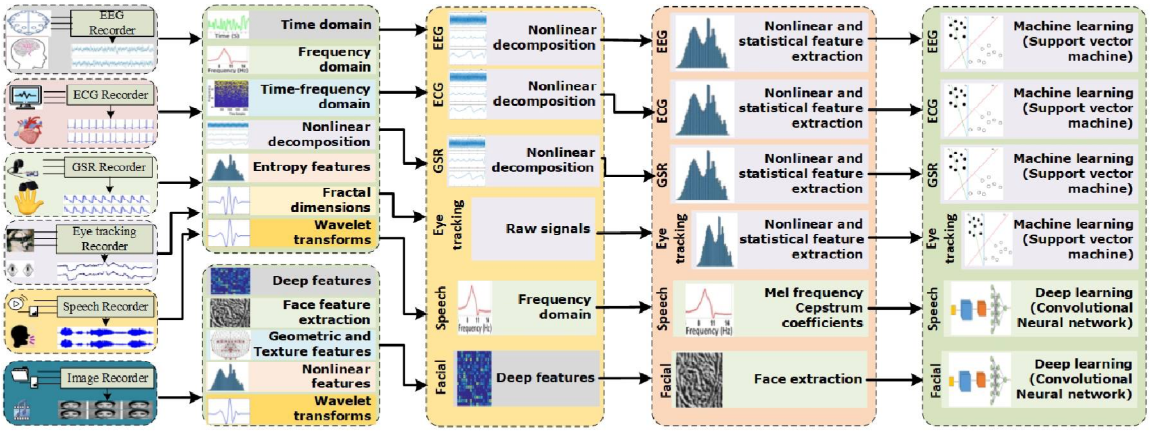


Distribution for emotion recognition studies using EEG signals



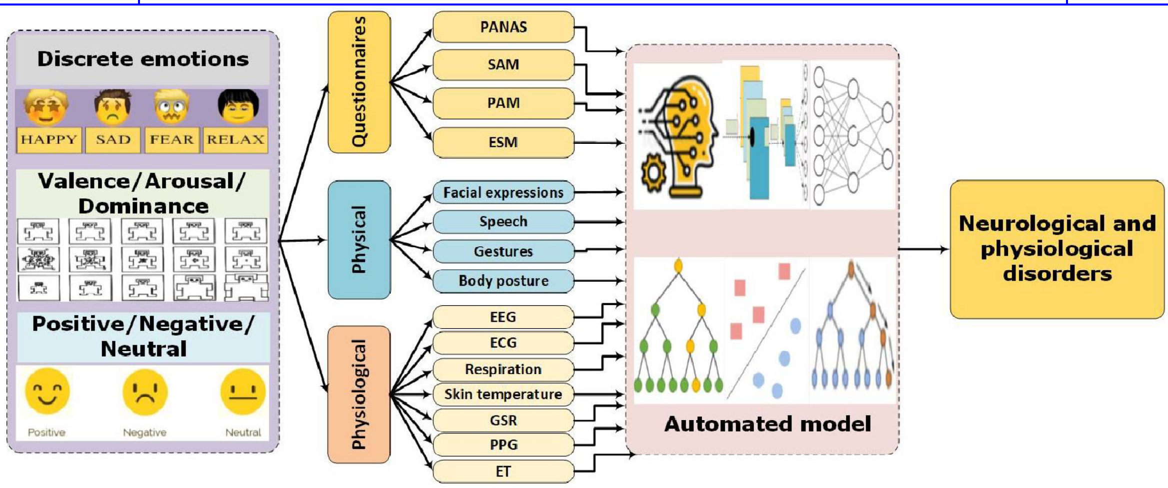
Distribution for emotion recognition studies using ECG signals

xAI.	Medical, Emotion	2023-161



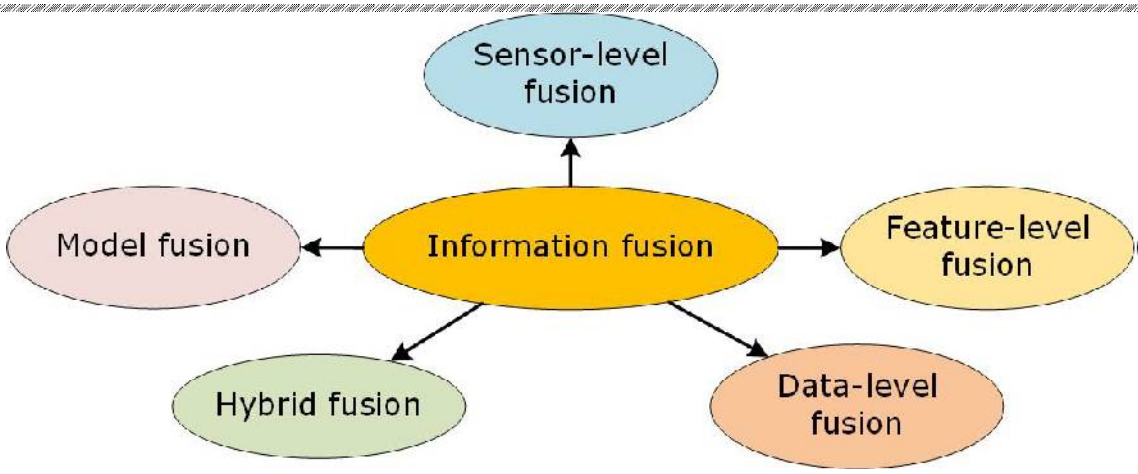
Graphical representation and summary of included modalities emotion recognition

xAI.	Medical, Emotion	2023-161



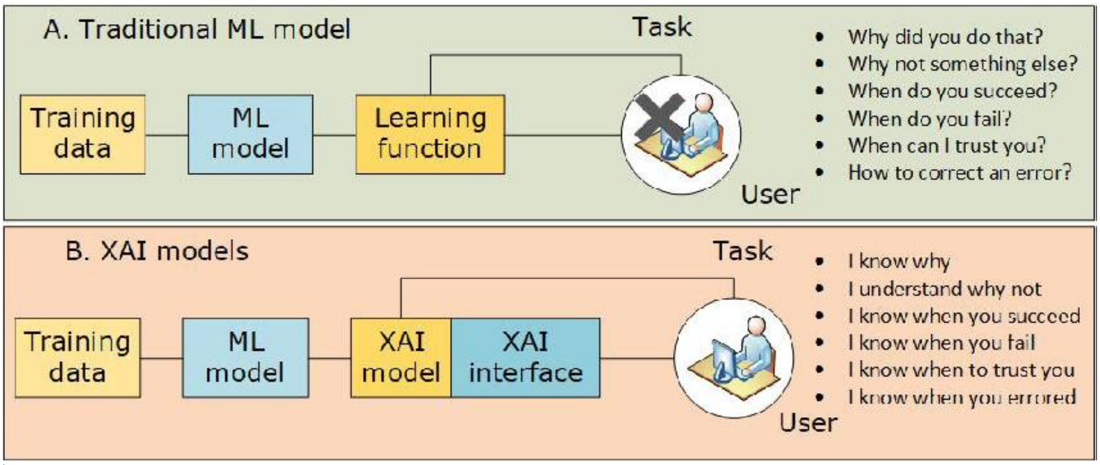
Emotion-based automated disorder detection system

xAI.	Medical, Emotion	2023-161



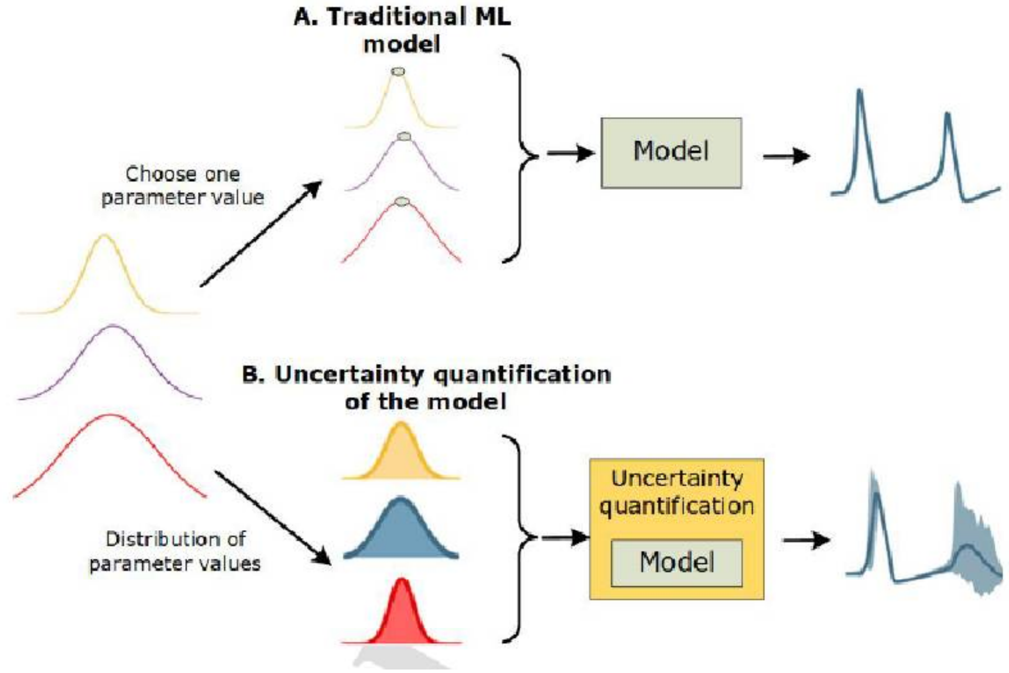
Taxonomy of information fusion

xAI.	Medical, Emotion	2023-161



- ✓ Illustrative representation of XAI model
- ✓ (A) Traditional ML model (B) XAI model with explanations

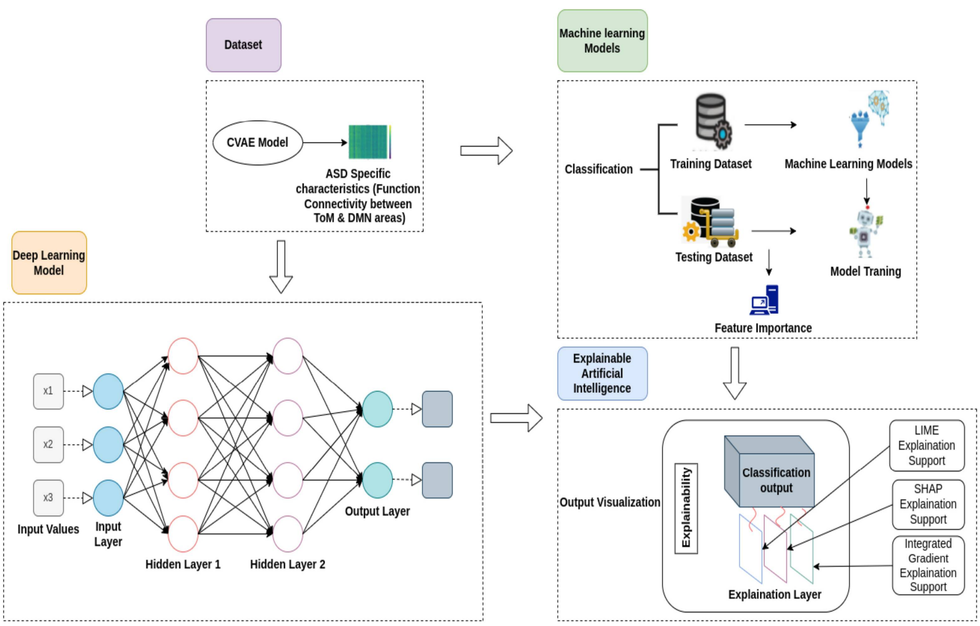
xAI.	Medical, Emotion	2023-161



- ✓ Uncertainty quantification of deterministic model
- ✓ (A) Traditional model with fixed parameter setting; (B) An UQ of the model with distributed parameter settings

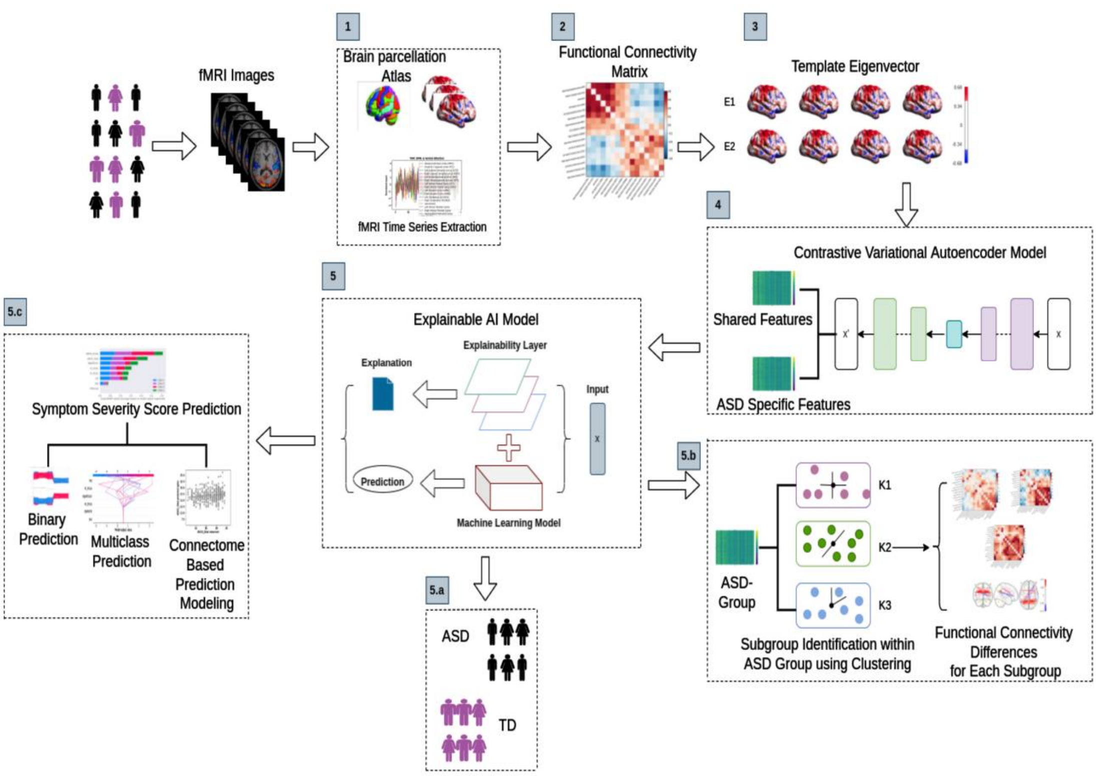
xAI.	Med	2023-137

Architecturefor ASD

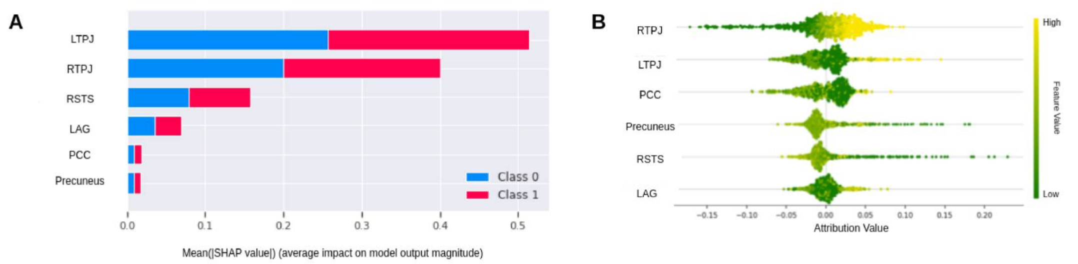


xAI.	Med	2023-137

ASD Classification

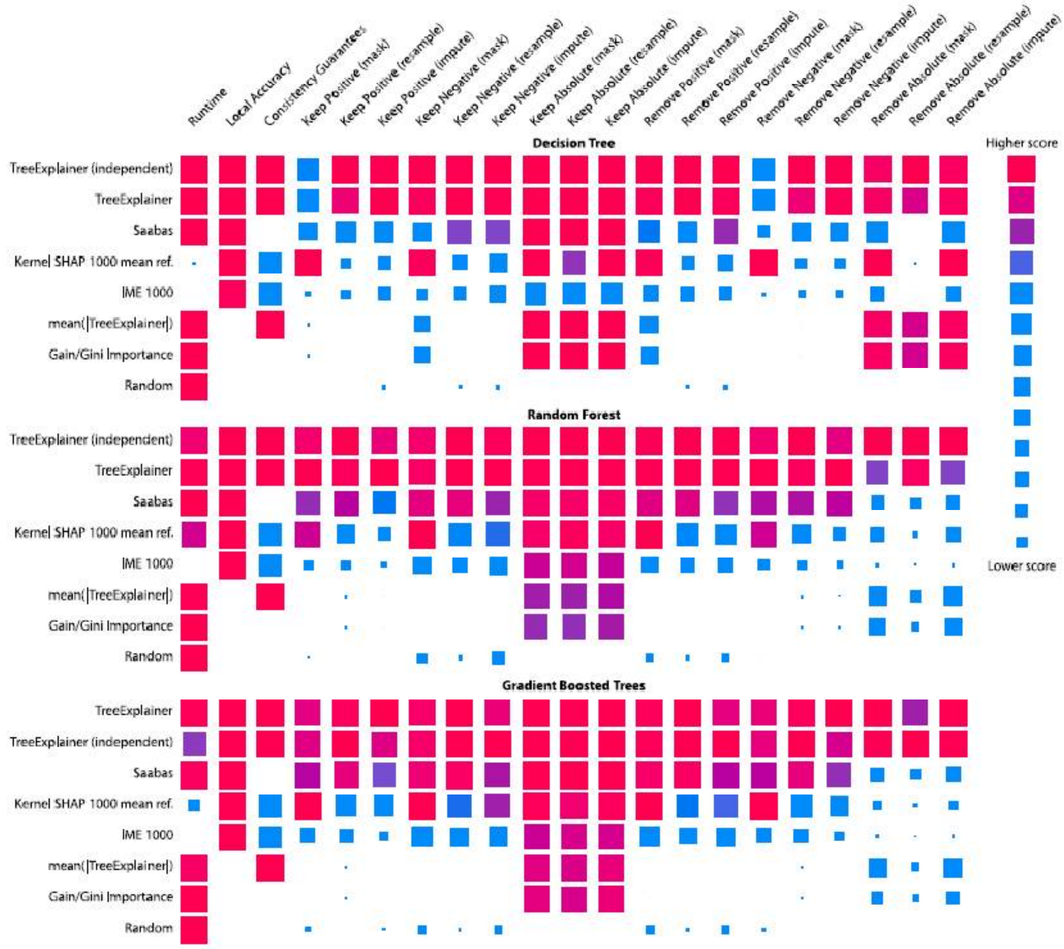


Importance of Features underlying
classification of ASD and TD



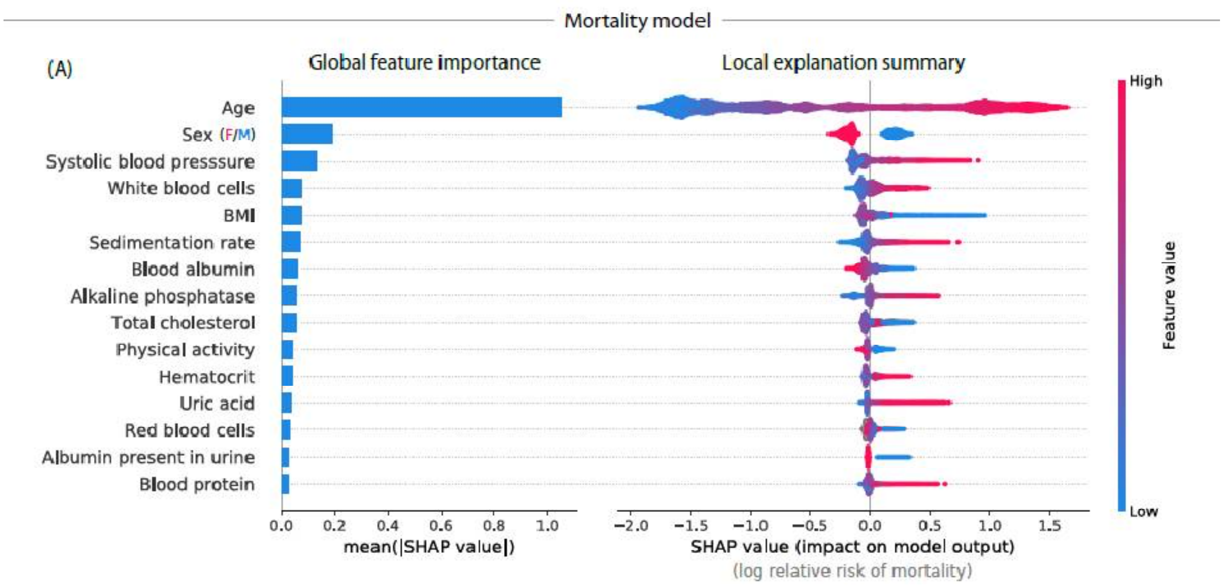
xAI.		2023-158

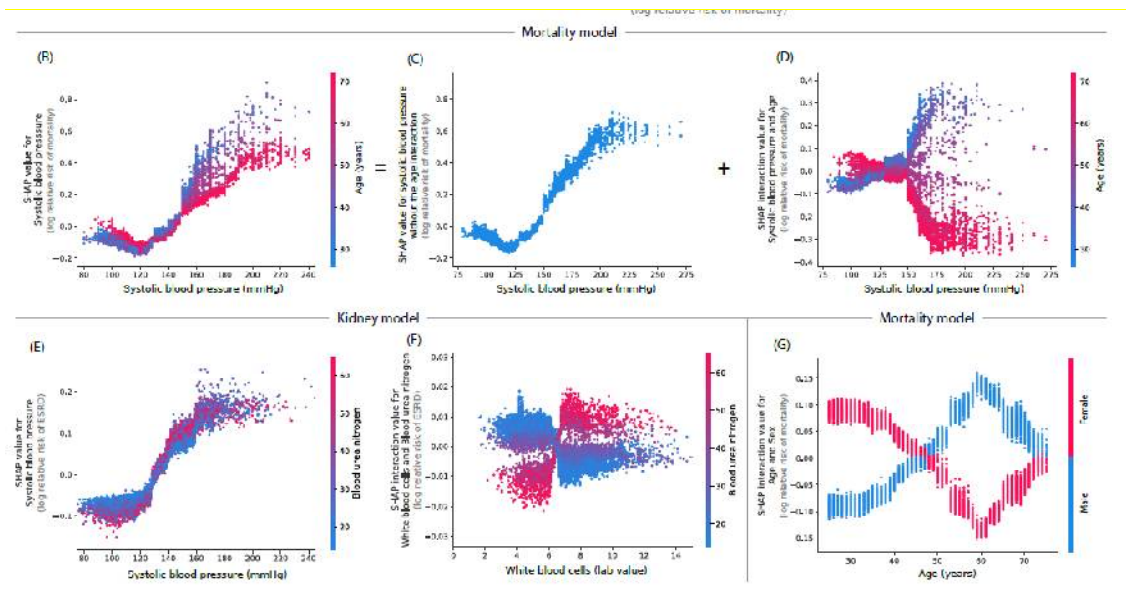
Kidney disease chronic
21 Evaluation metrics and 3 classification models



- ✓ TreeExplainer outperforms previous approaches
  - o Not only by having theoretical guarantees of consistency,
  - o But also across a large set of other metrics (Methods 11).

xAI.		2023-158
	Mortality model	



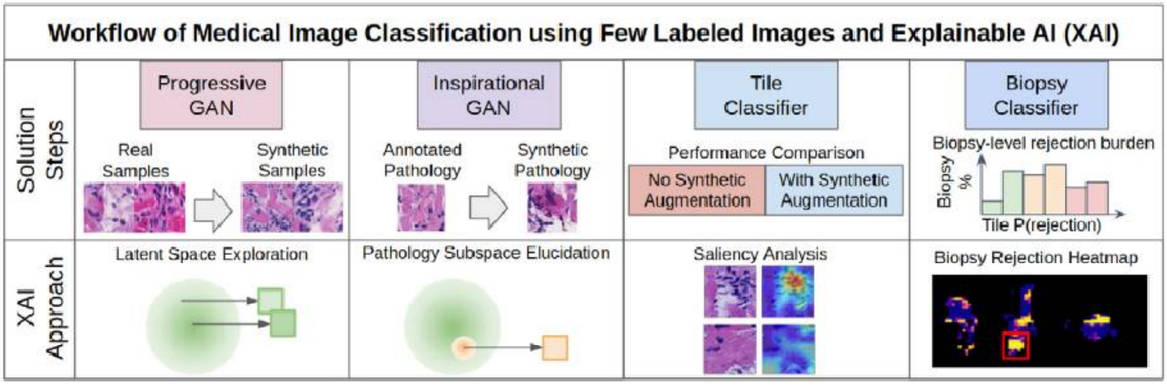


- ✓ By combining many local explanations, we can provide rich summaries of both anentire model and individual features.

(A) Bar chart (left) and SHAP summary plot (right) for a gradient boosted decision tree model trained on the mortality dataset. The long right tails in the summary plot are from rare but high-magnitude risk factors.
(B) SHAP dependence plot of systolic blood pressure vs. its SHAP value in the mortality model. A clear interaction effect with age is visible that increases the impact of early onset high blood pressure.
(C) Using SHAP interaction values we can remove the interaction effect of age from the model.
(D) Plotting just the interaction effect of systolic blood pressure with age shows how the effect of systolic blood pressure on mortality risk varies with age. Adding the y-values of C and D produces B.
(E) A dependence plot of systolic blood pressure vs. its SHAP value in the kidney model shows an increase in kidney disease risk at a systolic blood pressure of 125 (which parallels the increase in mortality risk).
(F) Plotting the SHAP interaction value of ‘white blood cells’ with ‘blood urea nitrogen’ shows that

high white blood cell counts increase the negative risk conferred by high blood urea nitrogen.
(G) Plottingthe SHAP interaction value of sex vs. age in the mortality model shows how the differential risk of men andwomen changes over their lifetimes"

xAI.	Medical	2023-039



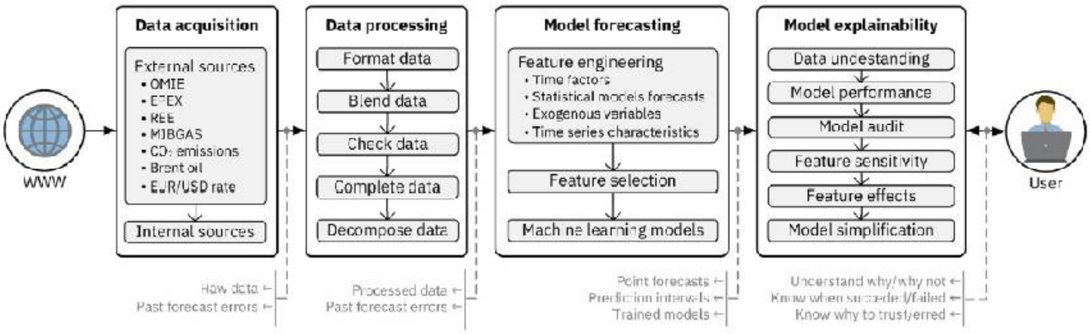
Overview of our approach to generate explainable clinical decision support tools for rare disease detection.

- ✓ "First, progressive and inspirational GANs were used to generate synthetic examples of normal and cellular rejection images, respectively, using limited expert-annotated examples. These generative models are interpreted by exploring low-dimensional projections of the input latent space.
- ✓ Trained two tile-level image classifiers to detect cellular rejection signs at high magnification. One classifier was trained using conventional augmentation alone (rotation and flipping), whereas the second was trained with conventional and synthetic image augmentation.
- ✓ Results were compared between the two approaches to demonstrate a significant performance boost with synthetic augmentation.
- ✓ Multiple saliency analysis methods (e.g. Grad-CAM++) were conducted to identify image regions used for successful and unsuccessful classification for both classifier models.
- ✓ The tile-level rejection probabilities were used to train a biopsy-level rejection classifier.
- ✓ In addition, biopsy-level rejection heatmaps were generated and compared with expert-labeled regions to support classifier labels"

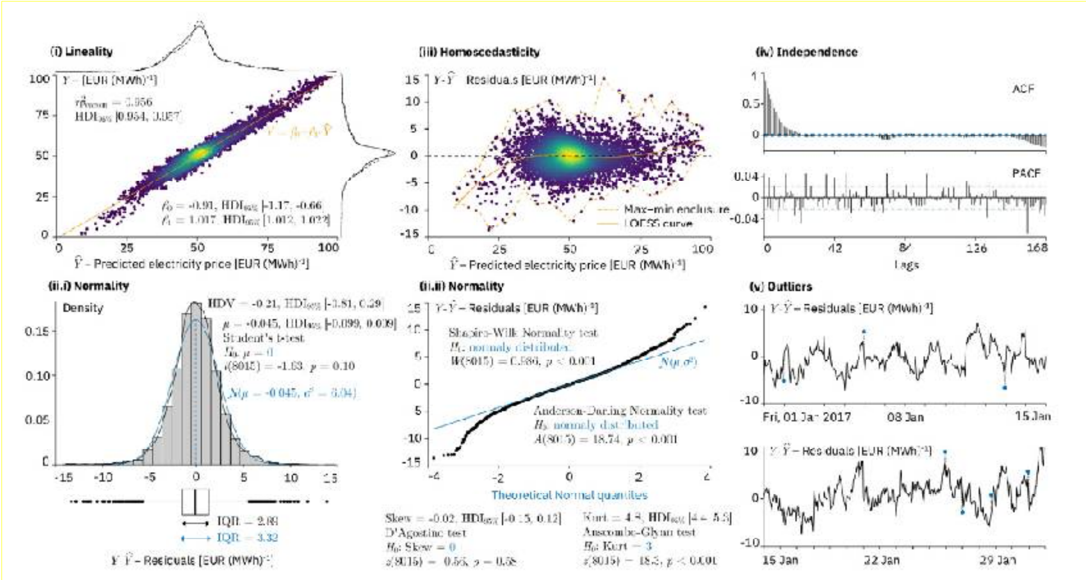
Electricity price forecast

xAI.	Electricity price	2022-163

Framework for **collaborative intelligence**
in forecasting day-ahead electricity price

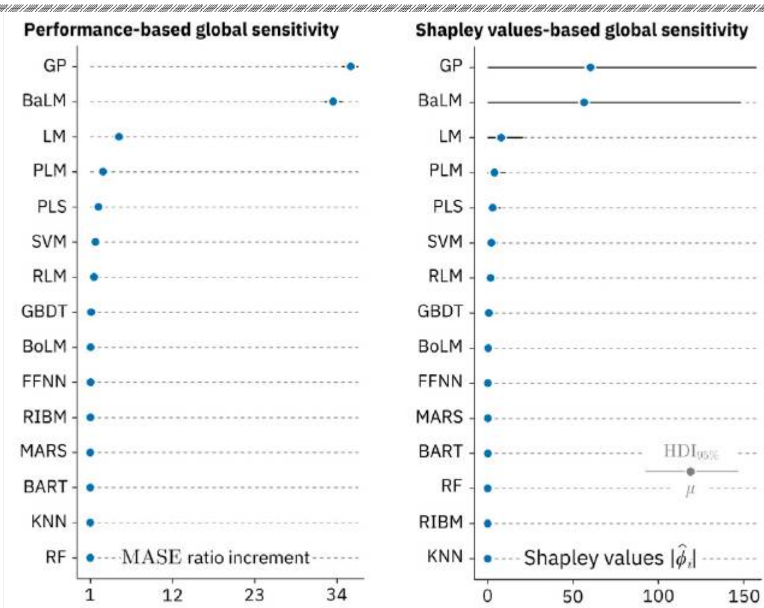


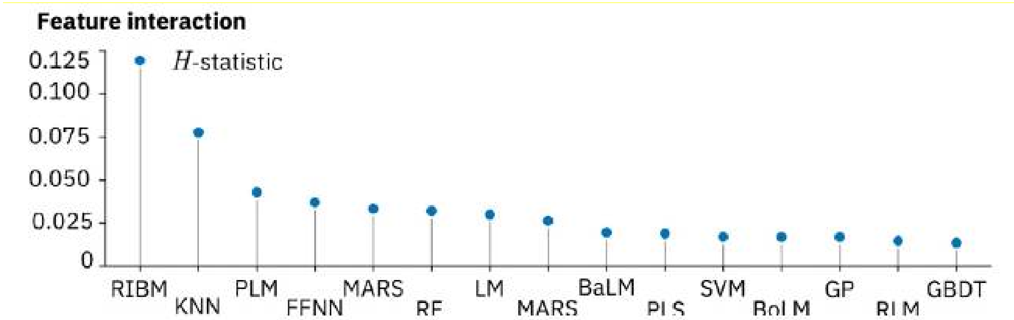
Visual representations of the Model Audit module

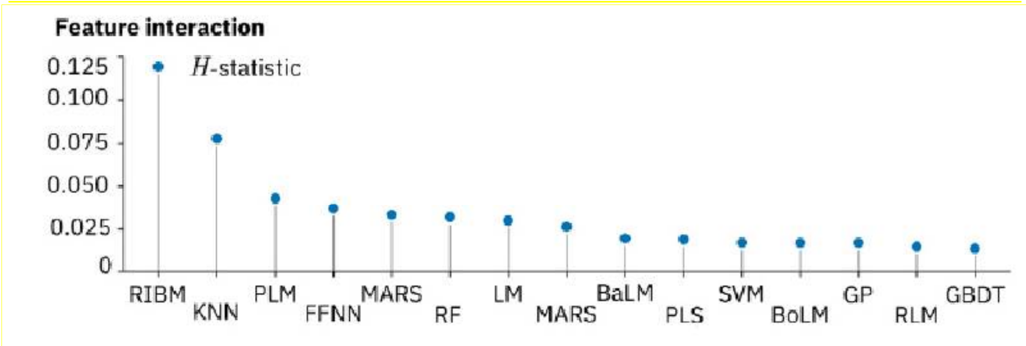


(i) Linearity inspection between observed and predicted electricity price.
(ii) Normality examination of the distribution of residuals by histogram (i) and quantile–quantile plot
(iii) Homoscedasticity audit of residuals.
(iv) Independence inspection of residuals.
(v) Outlier detection of residuals

Global sensitivity plots for the features of Level 2 of the stack ensemblearchitecture.


Performance-based global sensitivity
GP
BaLM
LM
PLM
PLS
SVM
RLM
GBDT
BoLM
FFNN
RIBM
MARS
BART
KNN
RF
MASE ratio increment
1
12
23
34
Shapley values-based global sensitivity
GP
BaLM
LM
PLM
PLS
SVM
RLM
GBDT
BoLM
FFNN
MARS
BART
RF
RIBM
KNN
HDI95%
μ
Shapley values $|\hat{\phi}_i|$
0
50
100
150


Feature interaction
H-statistic
0.125
0.100
0.075
0.050
0.025
0
RIBM
KNN
PLM
FFNN
MARS
RF
LM
MARS
BaLM
PLS
SVM
BoLM
GP
RLM
GBDT


Feature interaction
H-statistic
0.125
0.100
0.075
0.050
0.025
0
RIBM
KNN
PLM
FFNN
MARS
RF
LM
MARS
BaLM
PLS
SVM
BoLM
GP
RLM
GBDT

Features based on time series characteristics.

Nº	Description
1–5	Average, Median, Std. dev., Max., Min.
6	|Max. – Min.|
7–9	First, Second, Third value
10–12	Last, Second last, Third last value
13	Spectral Shannon entropy
14–15	Stability, Lumpiness
16–18	Max level shift, Max var shift, Max kl shift
19	Crossing points
20	Flat spots
21	Hurst
22–25	PACF features: (x, diff1, diff2, seas)-pacf5
26–27	Holt's linear trend method: α, β
28–34	STL features: nperiods, seasonal period and strength, trend, spike, linearity, peak

No	Description
35–37	Holt–Winter's seasonal method: α, β, γ
38–41	Heterogeneity: (ARCH, GARCH)–ACF, –R^2
42	Non linearity
43	ARCH statistic
44–47	Correlation: Embed2 , AC9, FirstMin, trev
48–49	Distri.: HistogramMode, OutlierInclude
50	Entropy: SampEn
51–52	Forecasting: LocalSimple, LoopLocalSimple
53	Non-linear time-series analysis: FluctAnal
54–55	Stationary: Std1thDer, SpreadRandomLocal
56	Symbolic transformations: MotifTwo
57	Others: Walker
58–66	ACF features: (e, x, diff1, diff2, seas)–acf1, (e, x, diff1, diff2)–acf10

List of statistical time series models whose forecasts are used as features.

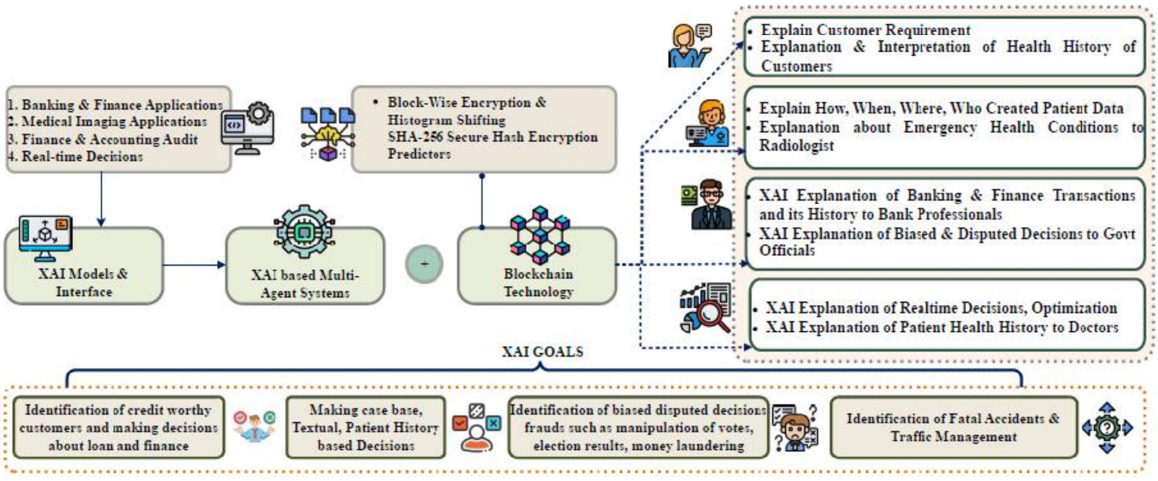
№	R-package::function	Description
1	fitAR::fitAR	AR(p) fitting
2	fGarch::garchFit	GARCH fitting
3	forecast::Arima	ARIMA(p, d, q)(P, D, Q)
4	forecast::dshw	Double-Seasonal Holt–Winters method
5	forecast::ets	Exponential smoothing state space model
6	forecast::(s)naive	(Seasonal) *naïve* model
7	forecast::nnetar	Feed-forward neural network with one hidden layer
8	forecast::tbats	TBATS model
9	forecast::thetaf	Theta method
10	forecTheta::dotm	Dynamic optimized Theta model
11	forecTheta::dstm	Dynamic standard Theta model
12	forecTheta::otm	Optimized Theta model
13	forecTheta::stheta	Standard Theta method
14	forecTheta::stm	Standard Theta model
15	glmnet::glmnet	Generalized linear model with lags
16	greybox::alm	Advanced linear model with lags
17	greybox::lmCombine	Linear model with combined lags
18	greybox::lmDynamic	Linear model with combined lags
19	greybox::stepwise	Linear model with stepwise selection of lags
20	MAPA::mapaest	Mutliple aggregation prediction algorithm
21	nnfor::elm	Extreme learning machine
22	nnfor::mlp	Multilayer perceptron
23	PSF::psf	Pattern sequence based forecasting
24	rugarch::arfimafit	ARFIMA fitting
25	rugarch::ugarchfit	GARCH fitting
26	smooth::ces	Complex exponential smoothing
27	smooth::es	Exponential smoothing in SSOE state-space form
28	smooth::gum	Generalized exponential smoothing
29	smooth::msarima	Multiple seasonal state-space ARIMA
30	smooth::sarima	State-space ARIMA
31	smooth::sma	Simple moving average in state space form
32	stats::HoltWinters	Holt–Winters filtering
33	TSPred::fittestMAS	Moving average smoothing
34	xgboost::(gblinear)	Regularized linear model with lags as regressors

Smart Cities

Role of xAI

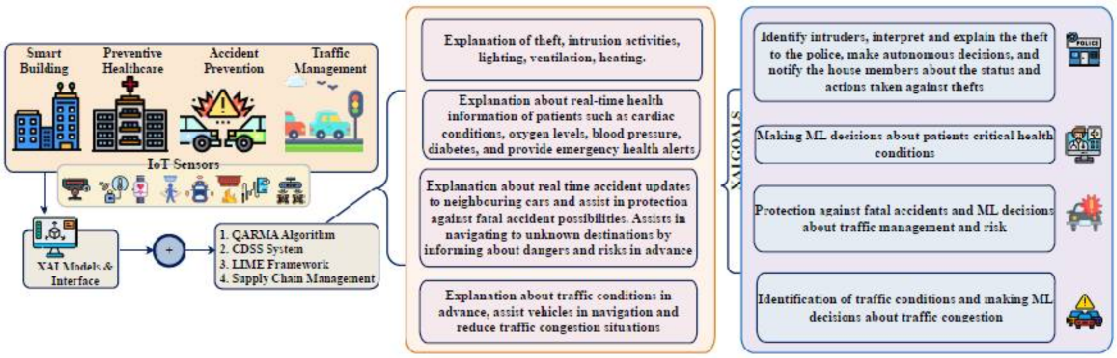
xAI. 2023-141

XAI with Blockchain for Smart Cities



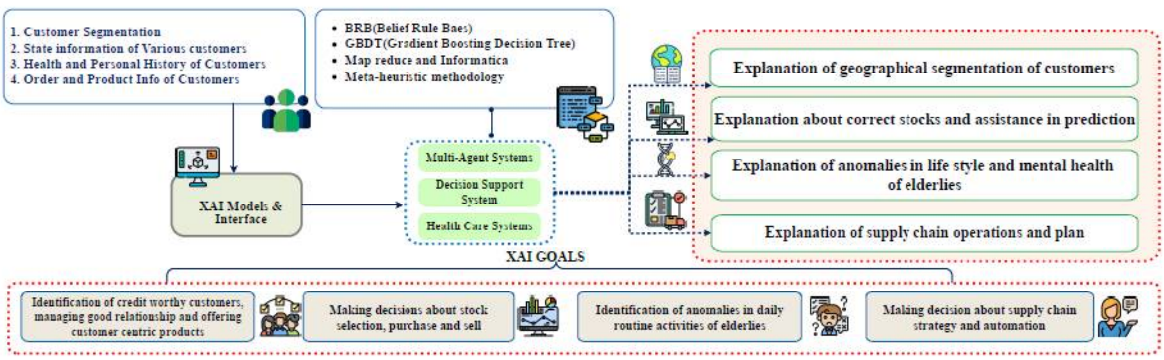
xAI. 2023-141

xAI with IoT for Smart Cities



xAI. 2023-141

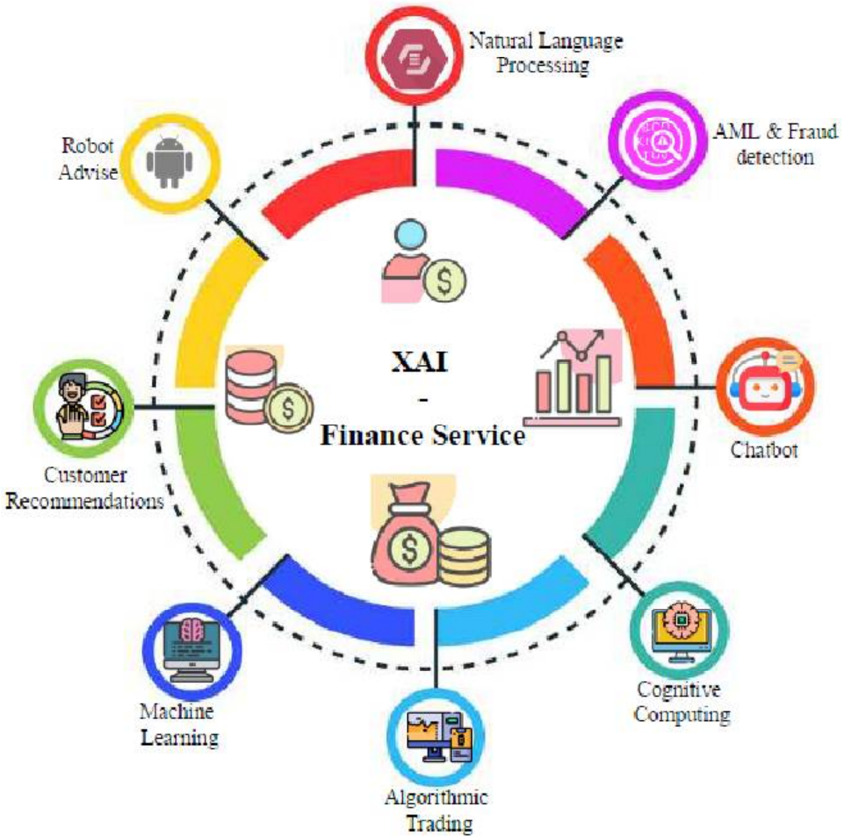
Usecase of XAI with Big Data for Smart Cities.



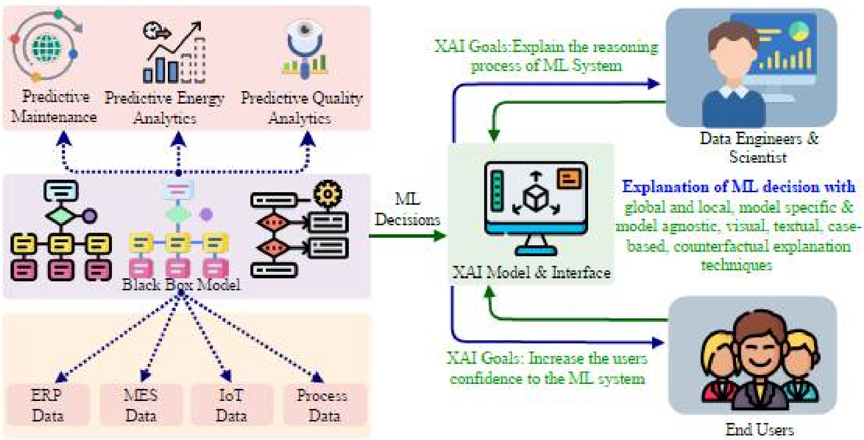
xAI Impact on
Financial Service

xAI.		2023-141

XAI in Financial Service



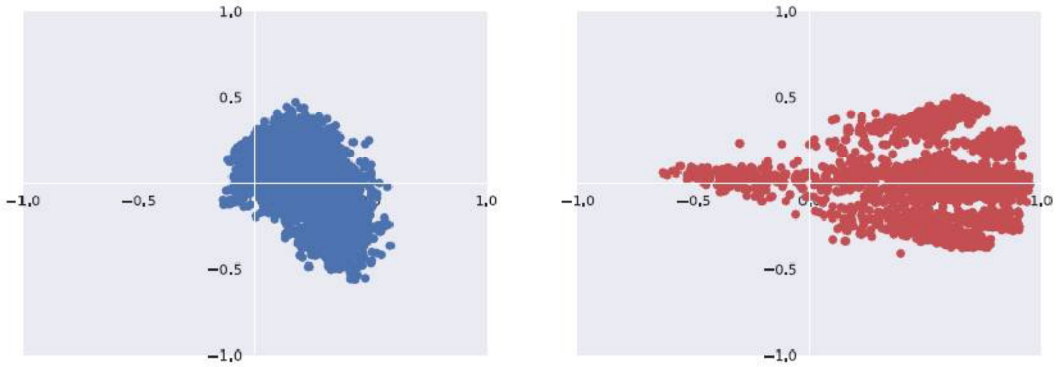
Role of XAI in Industry 4.0



Translation of one language -To-another

xAI.	Bilingual word embedding fusion	2023-146

Original English and Chinese embeddings

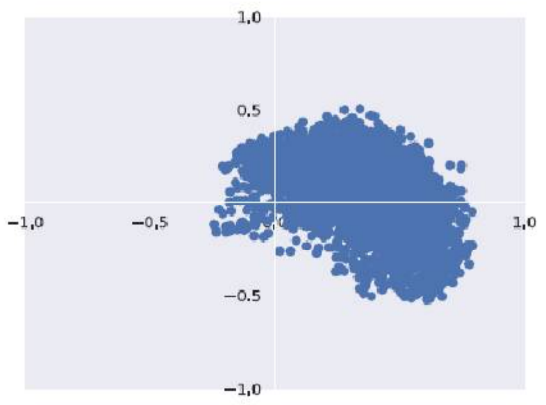


English ($s_1 = 14.19$ and $s_2 = 11.41$) Chinese ($s_1 = 42.25$ and $s_2 = 12.87$)

- ✓ $s1$ and $s2$ are the top two largest singular values

xAI.		2023-

English and Chinese embeddings after fusion

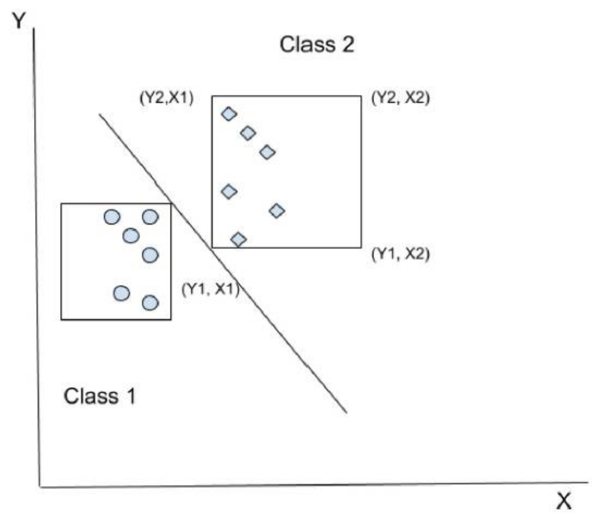


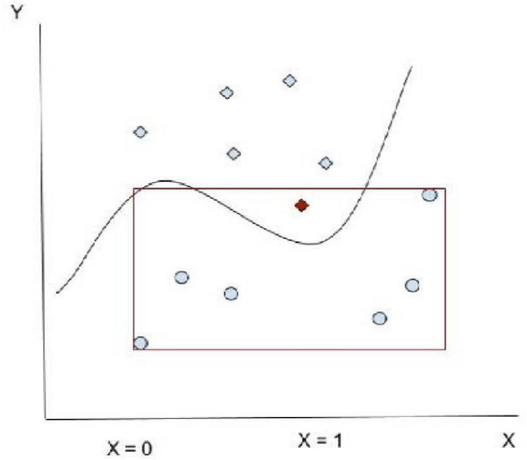
English ($s_1 = 25.82$ and $s_2 = 12.75$) Chinese ($s_1 = 26.06$ and $s_2 = 13.47$)

One-class SVM

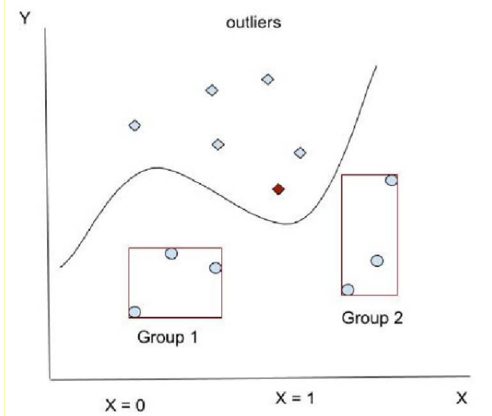
xAI.		2022-184

SVM with linear kernel classifying data points of two classes





- A hypercube generated using the farthest points leads to the wrong inclusion of data from another class

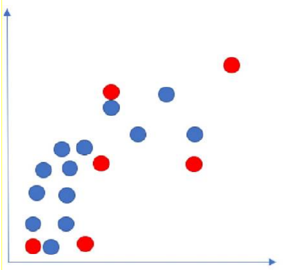


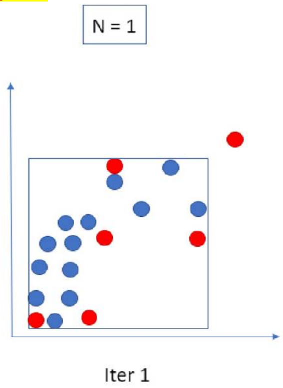
✓ Using more hypercubes avoids the aforementioned problem.

\+ Now there is no wrong inclusion of data points from another class

xAI.		2022-184

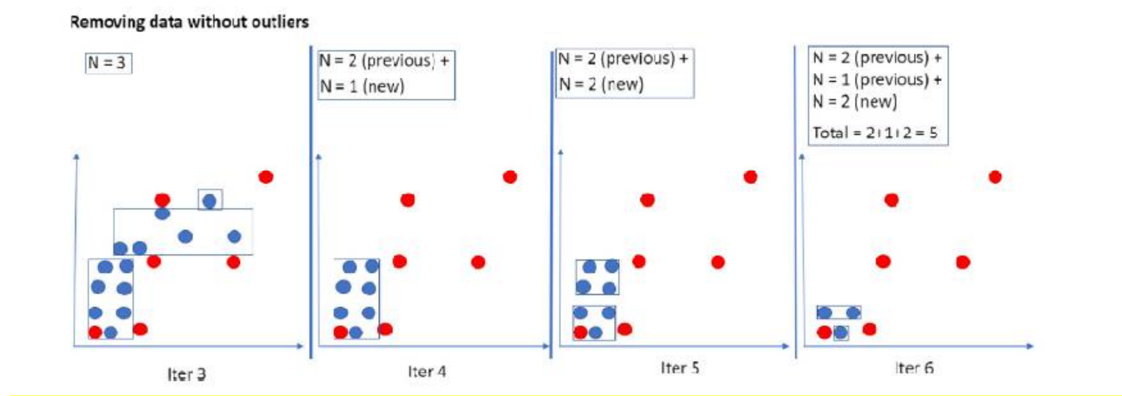
Clustering over a 2D space





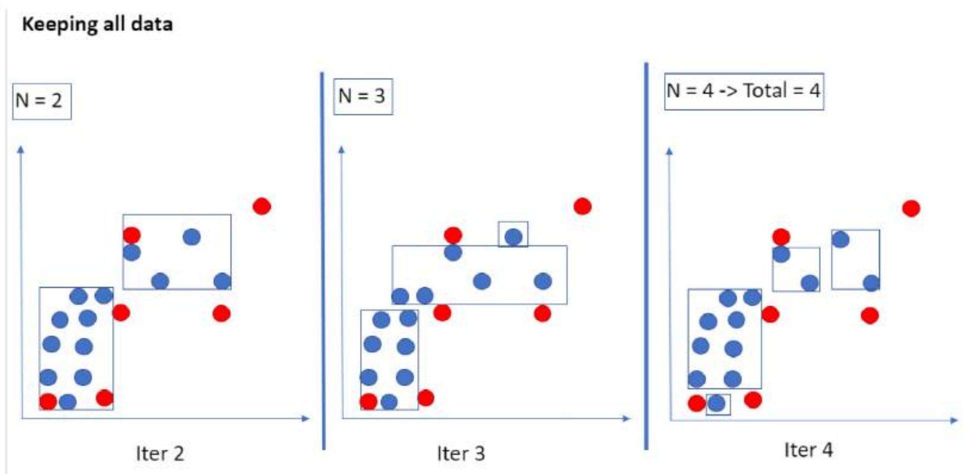
- o With one cluster over data points from one class (blue),
- o There are still others from the other class (red) inside the square

Applying Núñez et al. procedure

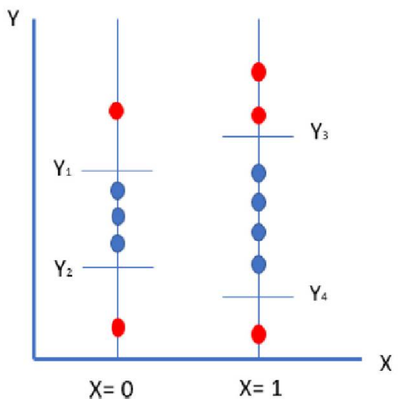


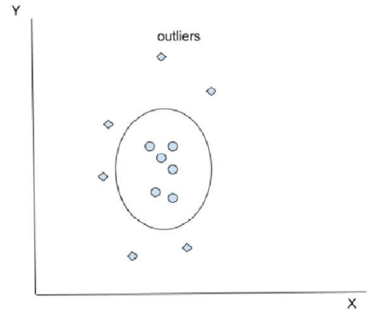
- ✓ Number of clusters keeps increasing until no points from the other class are inside
- ✓ Then that hypercube is translated into a rule

Keeping all data points in every iteration could lead to a reduced number of clusters

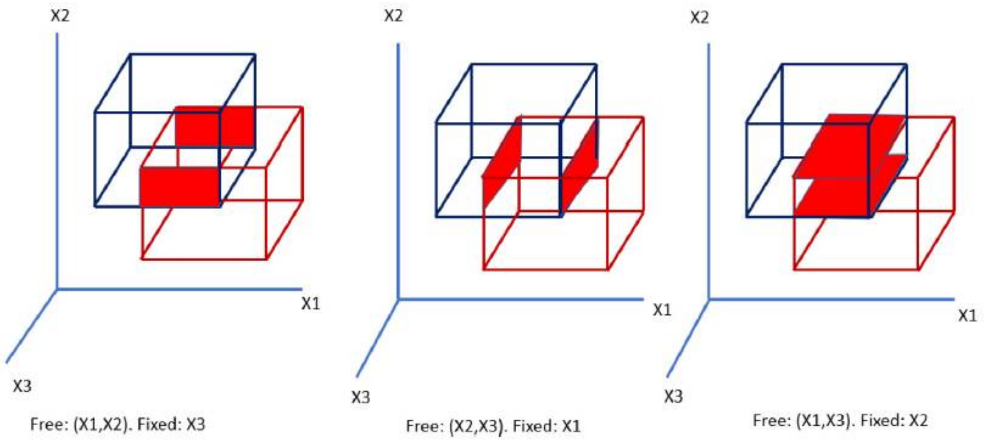


Rule extraction with a categorical variable	✓ With an RBF Kernel the correct hypercube will be the one that encloses the points that are not anomalies ✓ Because, OCSVM algorithm tries to enclose most of the points inside the decision frontier and leave anomalies outside





Overlapping between rules (hypercubes) approximated
using their 2D planes' area of intersection



Imbalanced-classes

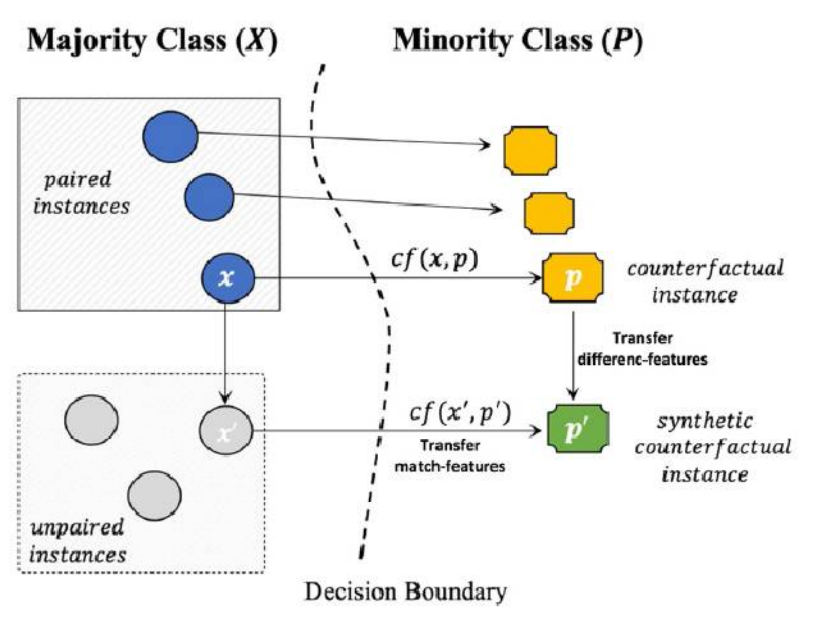
SMOTE
Synthetic Minority Over-Sampling Technique

xAI.		2022-069

Summary of SMOTE variants

Oversampling methods	References	Borderline regions	Selected regions	Majority usage	Noise removal	Clustering usage
SMOTE-Tomek	Batista et al. (2004)			✓	✓	
SMOTE-ENN	Batista et al. (2004)			✓	✓	
B-SMOTE	Han et al. (2005)	✓				
LLE-SMOTE	Wang et al. (2006)		✓			
ADASYN	He et al. (2008)		✓			
SL-SMOTE	Bunkhumpornpat et al. (2009)		✓	✓		
M-SMOTE	Hu et al. (2009)	✓			✓	
LN-SMOTE	Maciejewski and Stefanowski (2011)		✓			
SVM-SMOTE	Nguyen et al. (2011)	✓				
DBSCAN	Bunkhumpornpat et al. (2012)		✓			✓
SMOTE-RSB	Ramentol et al. (2012)			✓	✓	
NRSBoundary-SMOTE	Hu and Li (2013)	✓				
SDSMOTE	Li et al. (2014)	✓				
SMOTE-IPF	Sáez et al. (2015)	✓			✓	
GAS-SMOTE	Jiang et al. (2016)	✓				
SMOTE-D	Torres et al. (2016)	✓				
AND-SMOTE	Yun et al. (2016)		✓			
SOMO	Douzas and Bacao (2017)		✓			✓
CURE-SMOTE	Ma and Fan (2017)		✓			✓
k-Means SMOTE	Douzas et al. (2018)		✓			✓
Geometric-SMOTE	Douzas and Bacao (2019)		✓			
HCAB-SMOTE	Al Majzoub et al. (2020)	✓		✓	✓	
SWIM	Bellinger et al. (2020)			✓		

Counterfactual Augmentation (CFA)



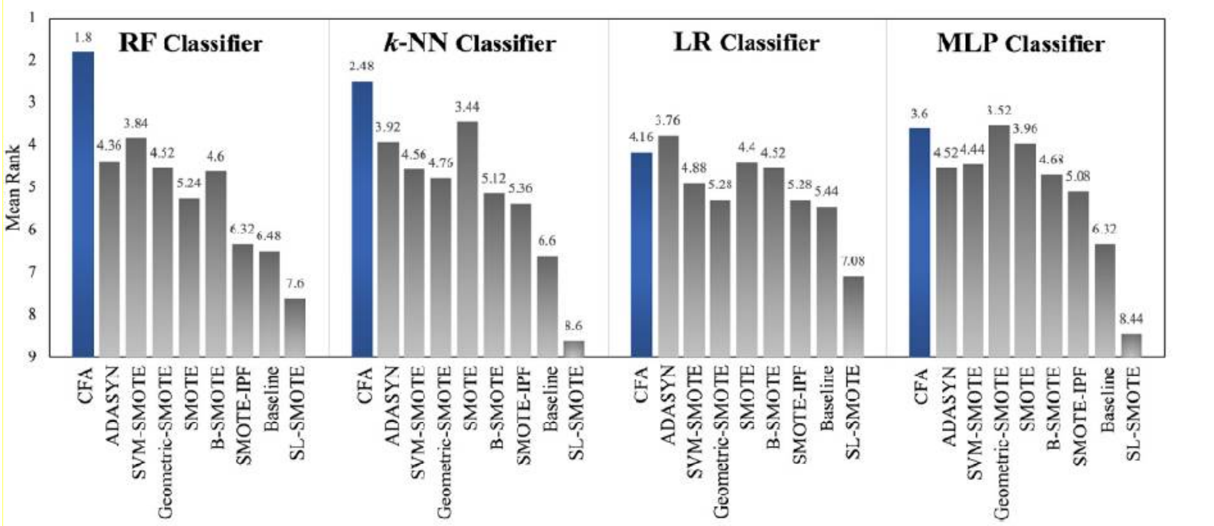
xAI.		2022-069

Typical Datasets & Imbalance Ratio (IR)

ID	Dataset	Features	Instances	Minority	Majority	IR
D1	Pima	9	768	268	500	1.86000
D2	Phoneme	6	5404	1586	3818	2.40000
D3	Vehicle	19	846	199	647	3.25000
D4	Abalone-9-vs-13	9	892	203	689	3.39000
D5	Yeast-3-vs-R	9	1484	163	1321	8.10000
D6	Ecoli-3-vs-R	8	336	35	301	8.60000
D7	Page-Blocks-0-vs-R	11	5472	559	4913	8.78000
D8	Yeast-0-3-5-9-vs-7–8	9	506	50	456	9.12000
D9	Abalone-9-vs-16	9	756	67	689	10.2800
D10	Glass-3-vs-R	10	214	17	197	11.5800
D11	WineQuality-Red-4-vs-5	12	734	53	681	12.8400
D12	Yeast-1-vs-7	9	459	30	429	14.3000
D13	Ecoli-4-vs-R	8	336	20	316	15.8000
D14	Abalone-13-vs-R	9	4177	203	3974	19.5700
D15	Abalone-9-vs-19	9	721	32	689	21.5300
D16	Abalone-9-vs-20	9	715	26	689	26.5000
D17	Yeast-4-vs-R	9	1484	51	1433	28.0900
D18	WineQuality-Red-6-vs-8	12	656	18	638	35.4400
D19	Abalone-17-vs-7-8-9-10	9	2338	58	2280	39.3100
D20	Yeast-6-vs-R	9	1484	35	1449	41.4000
D21	WineQuality-White-3-vs-7	12	900	20	880	44.0000
D22	WineQuality-White-3-9-vs-5	12	1482	25	1457	58.2800
D23	Poker-8-9-vs-6	11	1485	25	1460	58.4000
D24	Abalone-20-vs-8-9-10	9	1916	26	1890	72.6900
D25	Poker-8-9-vs-5	11	2075	25	2050	82.0000

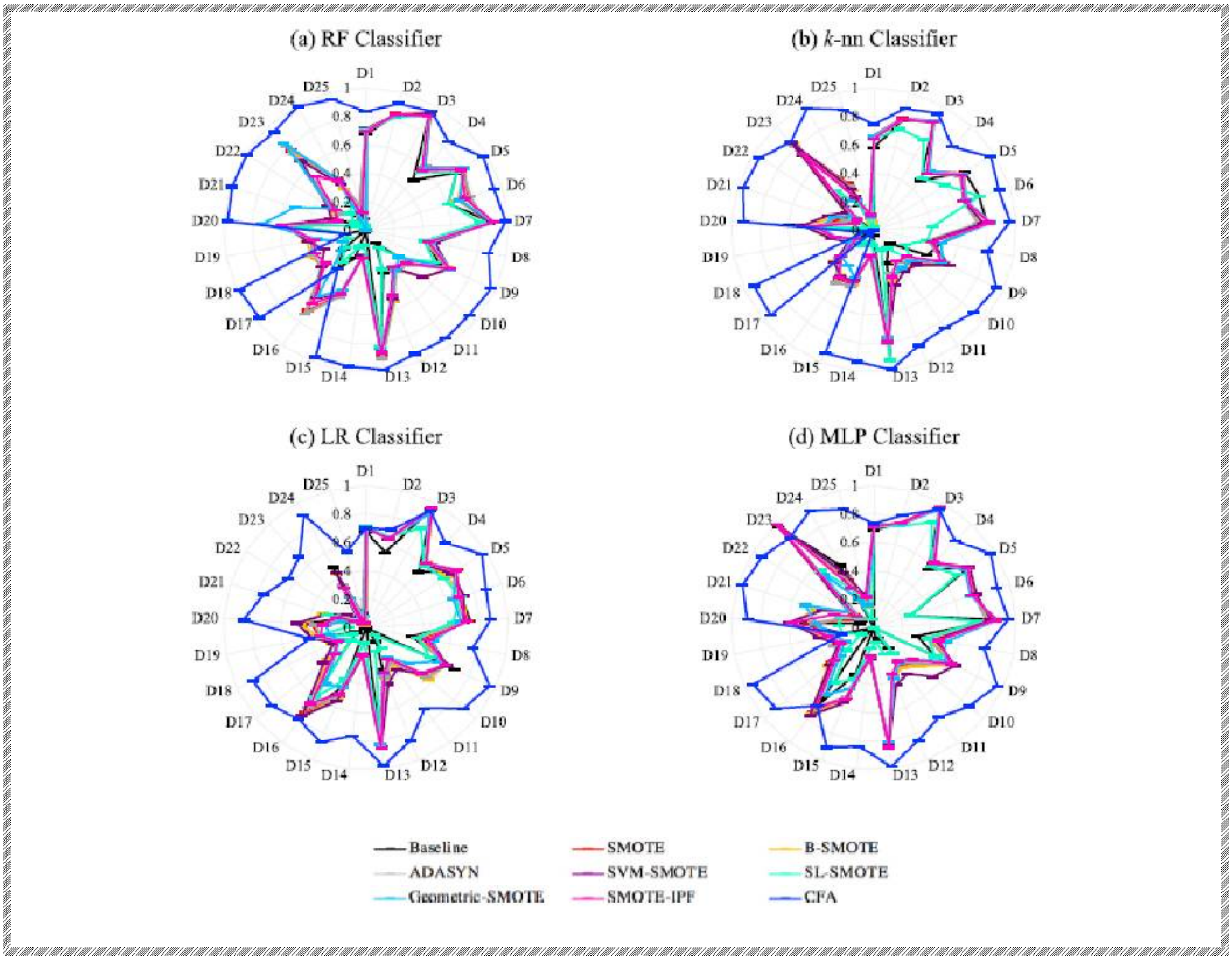
xAI. 2022-069

Mean ranking of each data augmentation method



xAI. 2022-

F1 values for the different conditions, across 25 datasets for the four classifiers
(a) RF classifier, (b) k-NN, (c) LR, and (d) MLP


(a) RF Classifier
(b) k-nn Classifier
(c) LR Classifier
(d) MLP Classifier
Baseline
SMOTE
B-SMOTE
ADASYN
SVM-SMOTE
SL-SMOTE
Geometric-SMOTE
SMOTE-IPF
CFA