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

ISSN: 2278-1862

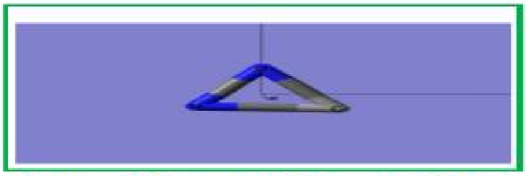
Journal of Applicable Chemistry

2024, 13 (2): 133-195

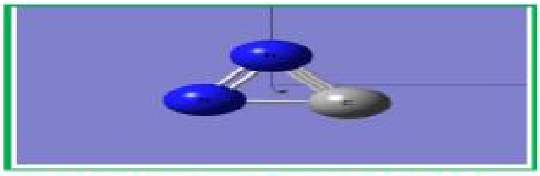
(International Peer Reviewed Journal)

New Chemistry News

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



New News of Chem (NNC)



ChemNewsNew (CNN)

CNN-59b—Fits (Figure Image TableScript...)Base (Bfits) xAI.Medicine (xAIM)-2024 Jan-Feb

Information Source	**sciencedirect.com;**	
S. Narasinga Rao M D **Associate Professor,** **Emergency Medicine dept.,** **Andhra Medical College,** **King George Hospital** **Visakhapatnam, A.P., 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: This news information document describes passive “Fits Base” with fields like Figure Image Table Script etc.AI based display methods (under development) will enable search, derivation of knowledge/information/intelligent sparkles. Thedisplay in intelligible format on screen and hard copy is another feature. This short baseis from medical research using xAI during the period JanandFeb of 2024. The records are picked upfrom standard abstract bases and full-on-line journals.

The diagnosis of critical diseases viz. pulmonary (covid-19), heart, brain (Alzheimer), diabetes, skin and. Bones are presented.

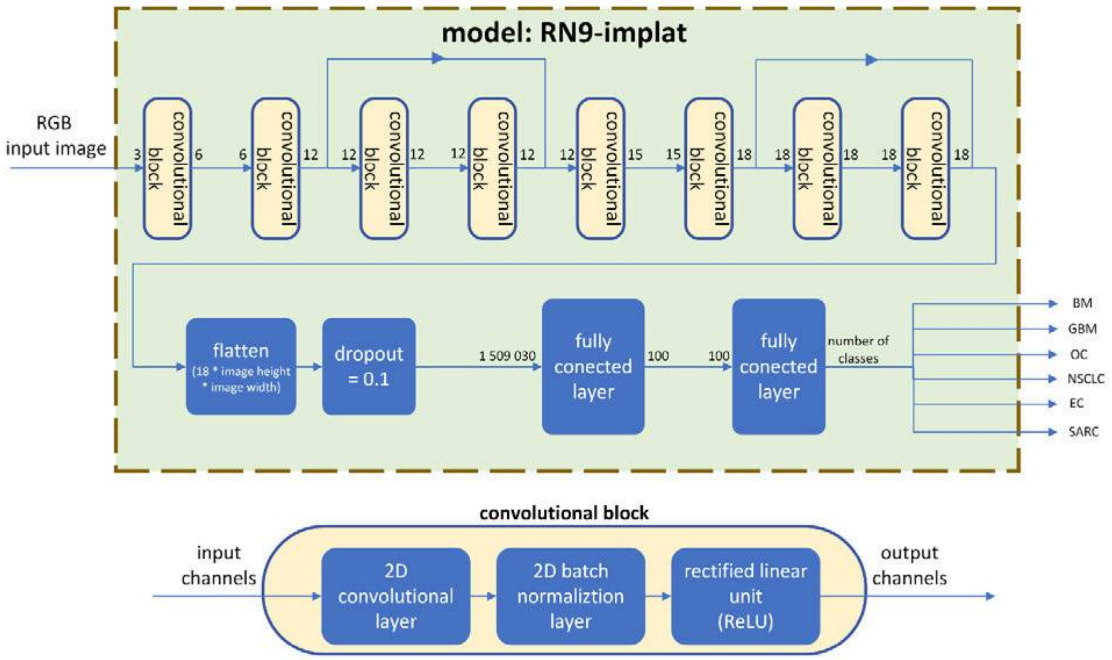
The computational methods used here are Machine learning (XGBOOST, Ensembles), RN9, fuzzy logic, Pretrained-Deep-NNs (YOLO, VGG 16/19) ,xAI stubs, Capsule Nets, eXplainable-Caps Nets etc. The xAI-probes employed in these studies includeShapley; LIME; CAM; Grad-CAM; Integrated gradients; Class Activation Maps; tSNE plot etc. These state-of-knowledge computational tools will pave toeXplainable/ interpretable/ Responsible/ Trustworthy AI products in application fields in the coming years.

***Keywords*:**eXplainable AI (xAI);covid-19, cardiac-diseases, Alzheimer, diabetes, skin problems. Orthopedic disorders.
Fits Base“([Figure, Fact, False], [Image; Information], [Table; Tensor; Truth], [Script ; Sound; Science]…) Base”
CNN : [C [Computations; Computer; Chemistry] NN [New News; News New; Neural Nets; Nature News; News of Nature;]]

The number refers to ref.No in CNN-59(a)

Architectures & Frames of methods

06



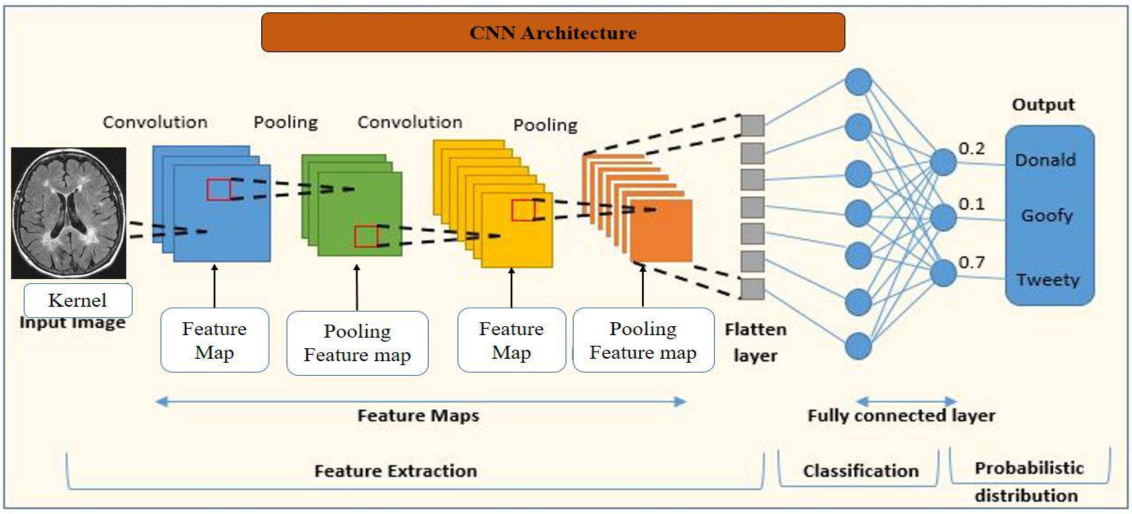
Model RN9-implat based on RN9 backbone.

- ✓ It comprises eight convolutional blocks for feature extraction and
- ✓ two fully connected layers responsible for classification tasks.
- ✓ A dropout layer was added to prevent overfitting.
  - o Numbers depict the amount of input and output channels foreach layer.

06

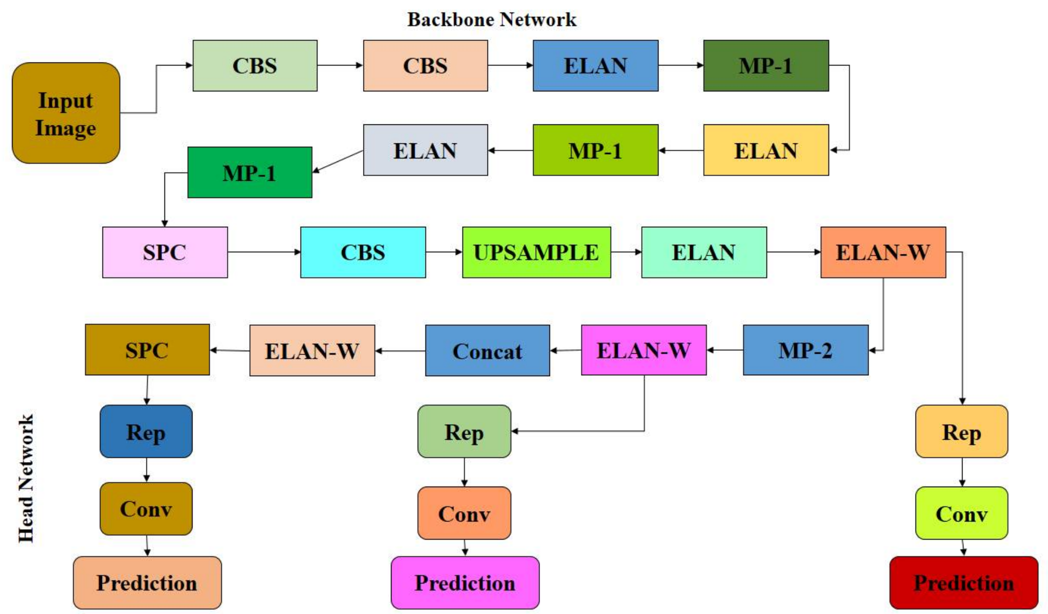
Methods	Approach	Finding	Dataset Used	Accuracy (%)	Class	Limitation
Convolutional Neural Network	Multi-class Classification	Improved tumor detection and classification	BRATS	90.5	Glioma, Meningioma, No Tumor, Pituitary	Limited by small dataset size, may struggle with rare tumor types.
Transfer Learning	Fine-tuning pre-trained models	Enhanced performance in low-data scenarios	MICCAI BraTS Challenge 2019 Training Data	88.2	Glioma, LGG	Dependency on the quality and representativeness of pre-trained models.
Recurrent Neural Network	Temporal sequence analysis	Improved temporal understanding of tumor growth	Hospital-based proprietary dataset	87.0	Glioma, Meningioma	Computationally intensive, limited scalability.
Ensemble Methods	Integration of multiple models	Increased robustness and generalization	TCGA Glioblastoma Multiforme Dataset	92.3	Glioblastoma Multiforme, Low-Grade Glioma	Complexity in model integration and interpretability.
3D Convolutional Networks	Volumetric image analysis	Improved spatial representation in brain tumor images	ISLES - Ischemic Stroke Lesion Segmentation	86.7	Stroke Lesions	Higher computational requirements, longer training times.
Capsule Networks	Hierarchical feature extraction	Enhanced feature learning for complex patterns	Figshare Dataset	89.6	Glioma, Meningioma, No Tumor	Limited interpretability of capsule networks.
Autoencoders	Unsupervised feature learning	Improved representation of latent features	RSNA Brain CT Hemorrhage Dataset	91.8	Intracranial Hemorrhage	Sensitive to noise in input data, may require careful preprocessing.

07



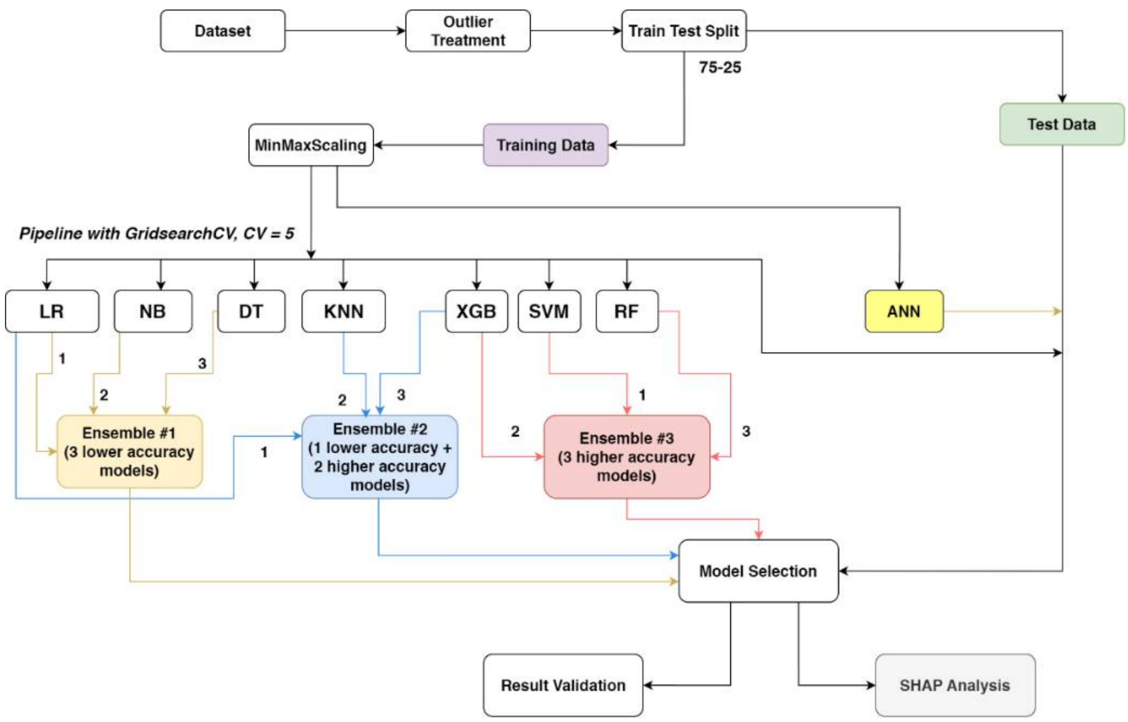
CNN Architecture

07



YoLO Architecture

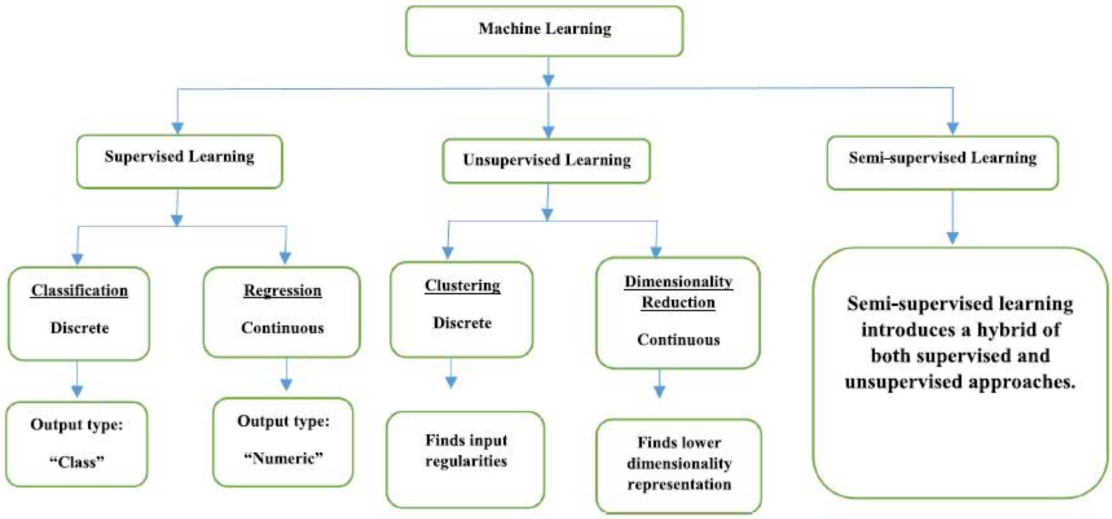
08



Methodology

ML methods

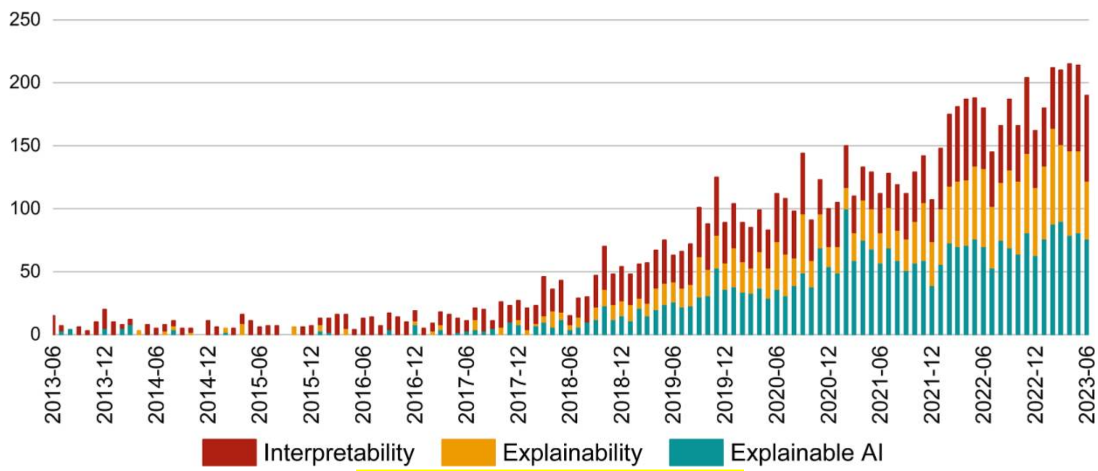
16



Overview of ML categories (Ikerionwu, et al., 2022)

Fuzzy methods

17



Google Trends data for the terms
"interpretability," "explainability," and "explainable AI"

- ✓ worldwide over the past 10 years, revealing an increasing in search interest since 2017

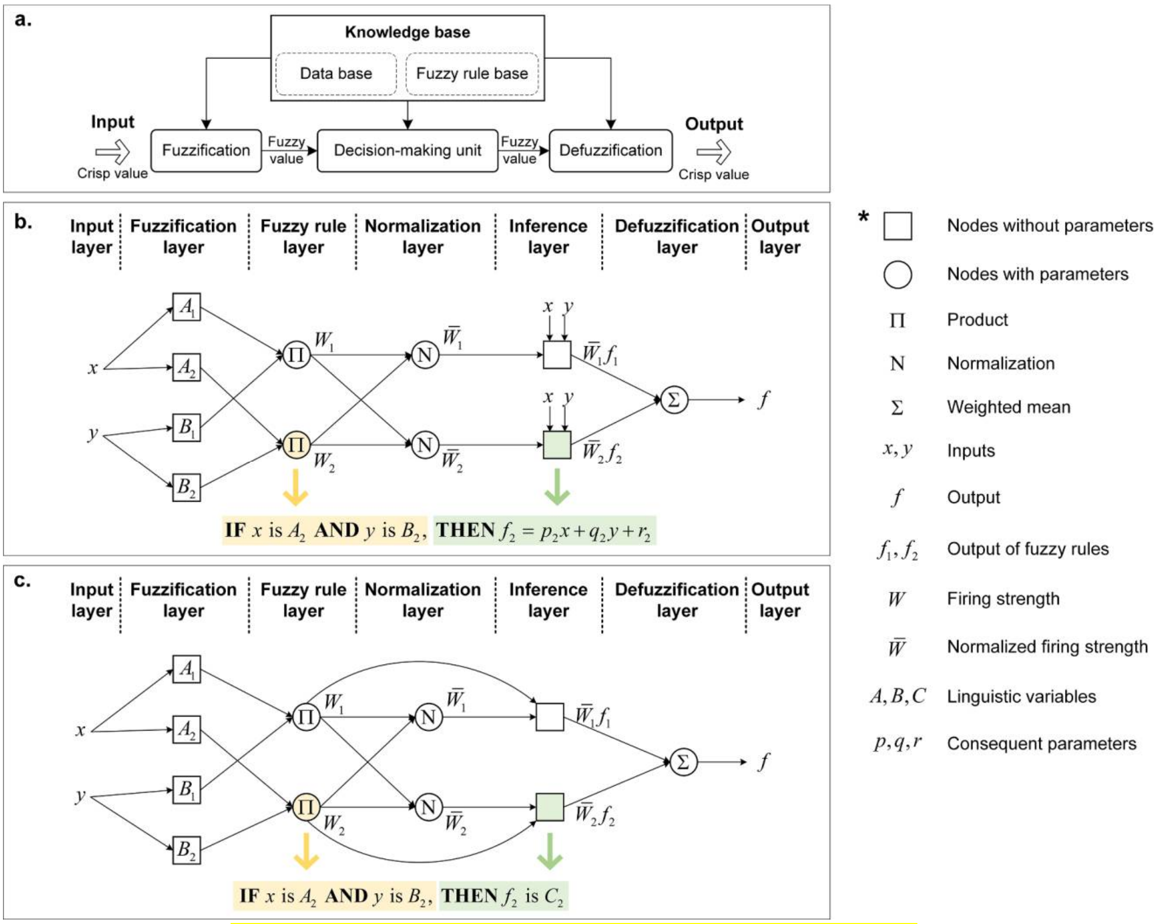
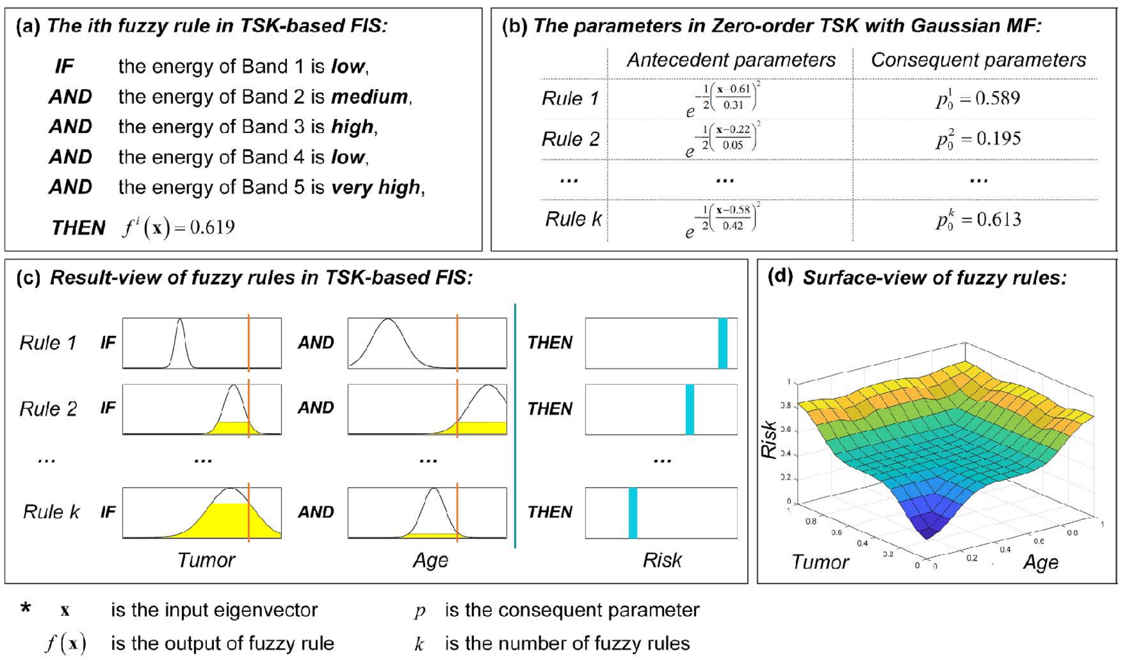


Illustration of the basic structures of an FIS and an ANFIS.

(a) Basic structure of an FIS.

(b) Network structure of a TSK-based ANFIS.

(c)Network structure of a Mamdani-based ANFIS



Presentations of fuzzy rules.

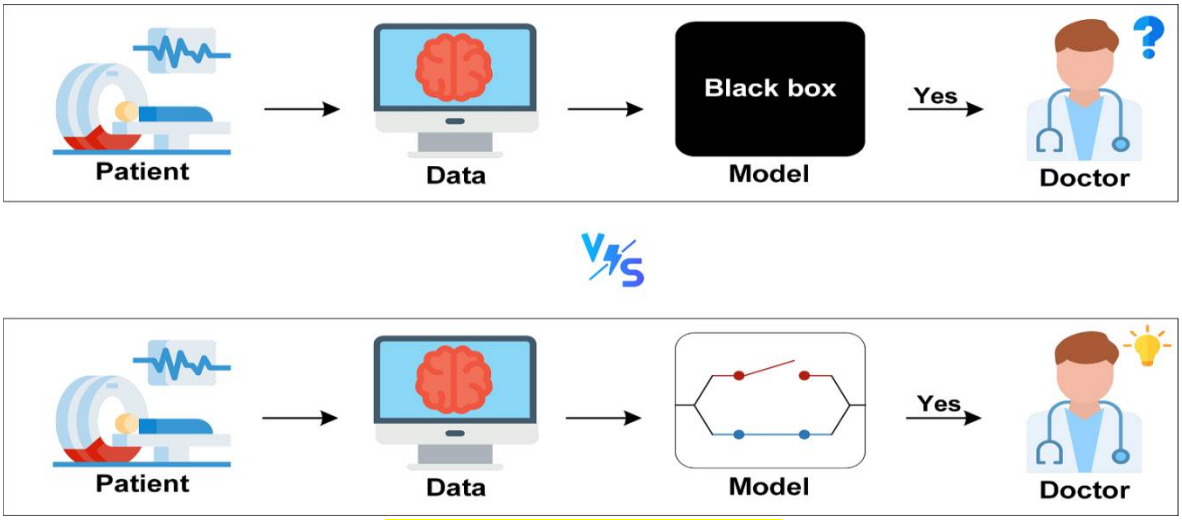
(a) An instance of fuzzy rules in the form of IF–THEN generated by the TSK-based FIS.

(b) An instance ofparameters in the zero-order TSK with a Gaussian membership function.

(c) An instance of a result-view of fuzzy rules.

(d) An instance of a surface-viewof fuzzy rules. The IF–THEN form and the visualized presentation of fuzzy rules are respectively the most understandable presentation for endusersand researchers in the fuzzy logic field

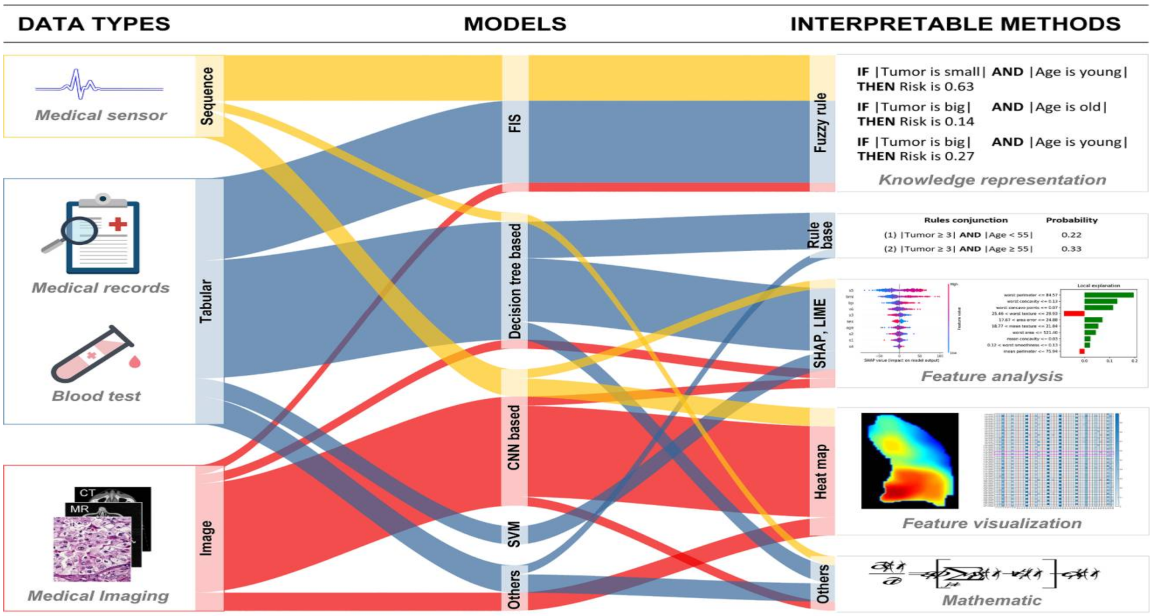
17



Computer-aided diagnosis system.

- ✓ Compared with a decision made using a black box model
- + Decision made using aninterpretable model is more understandable for doctors.

17



XAI techniques used in computer-aided disease diagnosis scenarios based on the analysis of 50 relevant literature reports

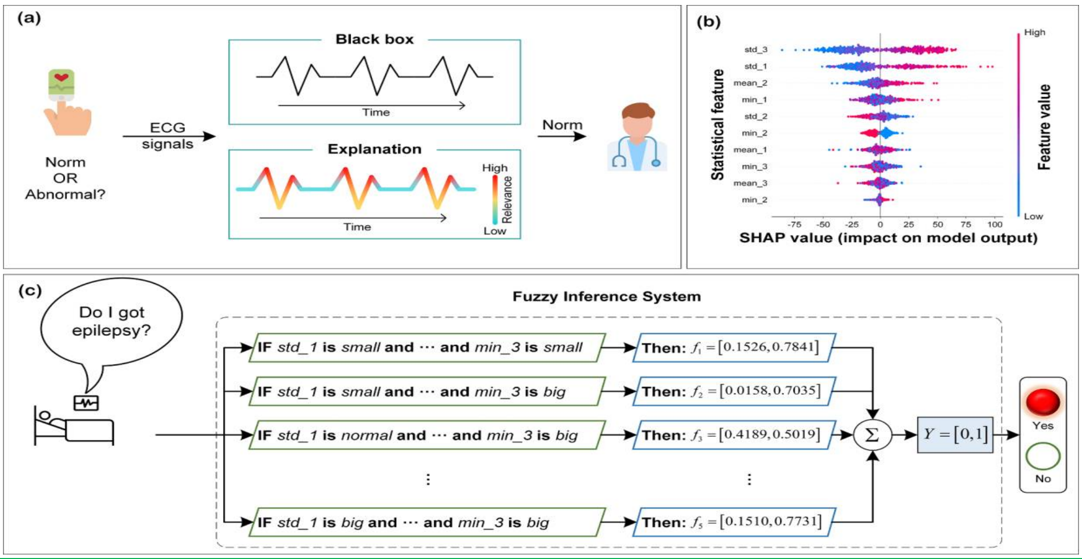
- ✓ The line width in the parallel set figure is determined by the number of relevant sources.
- ✓ The examples of data types and interpretability methods mentioned are not limited to those presented on the sides of the parallel set plot

17

Brief description of other interpretability methods.

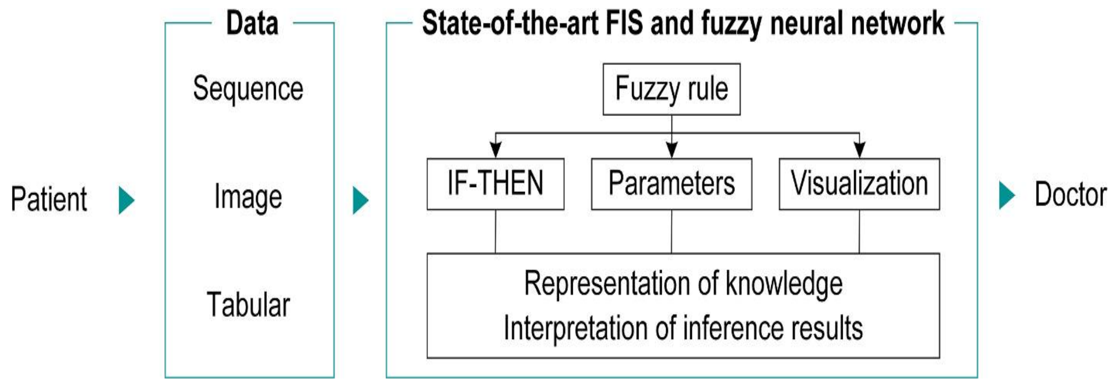
Interpretability methods	Description
Rule base	Rule base can be generated by decision tree-based model, including Decision Tree, Random Forest, XGBoost etc. These rules describe the conditions that lead to specific decisions, making the model easily interpretable. These models provide insights into the importance of each feature in the decision-making process, and it can also be well integrated with the SHAP principle. In the view of this, they are often used together. However, its biggest difference from fuzzy rules is that it does not include fuzzy linguistic variables, instead it relies entirely on crisp values, as shown in the panel of Fuzzy rule and Rule base in Fig. 5.
SHAP	SHAP is a game-theoretic approach that provides a unified framework for explaining the output of any machine learning model. It is based on concepts from cooperative game theory, specifically Shapley values, which allocate the contribution of each feature toward the prediction outcome. SHAP values represent the impact of each feature on the predicted outcome for a specific instance. These values enable us to understand the importance and influence of features in the model's output. An example is shown in the feature analysis panel of Fig. 5.
LIME	LIME is a technique for explaining the predictions of any black-box machine learning model. It aims to provide local and interpretable explanations by approximating the behavior of the model around specific instances. By examining the coefficients of the approximated model, LIME identifies which features were the most influential in influencing the prediction for that particular instance. These explanations help users understand the model's decision-making process at an individual instance level, thus increasing transparency and trustworthiness. An example is shown in the feature analysis panel of Fig. 5.
Heat map	A heatmap is a visualization technique used to represent the importance or relevance of features in a model. The color gradient in the heatmap helps identify patterns and correlations between features and instances. A higher intensity or a distinct color in a cell or pixel signifies a stronger influence of that feature on the model's decision, while lower intensity or a different color suggests a relatively lesser impact, as shown in the heat map panel in Fig. 5.

17



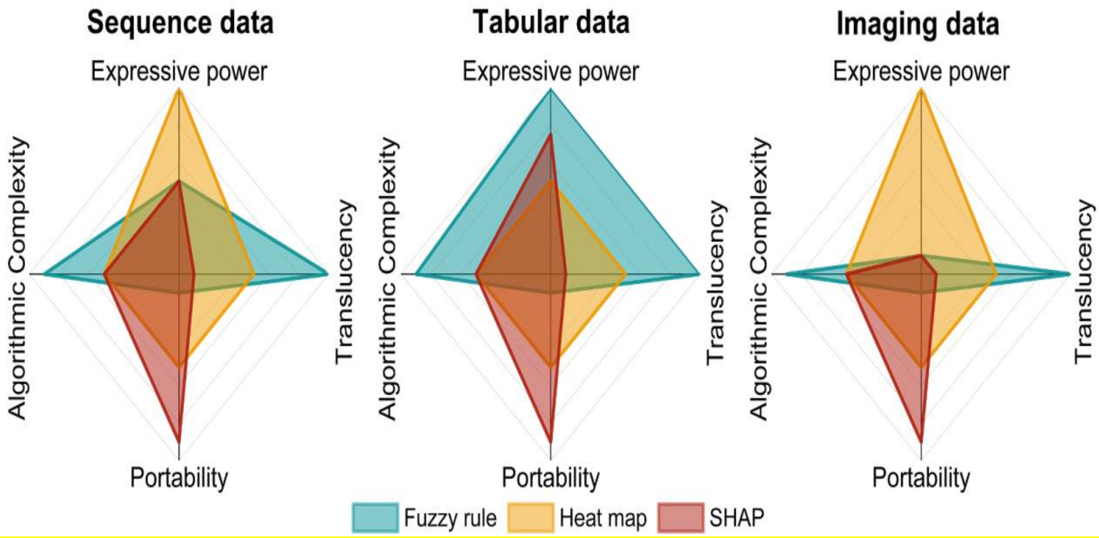
Explanation methods used in disease diagnosis with sequence data.
(a) A heat map used as an explanation tohighlight fragments with diverse relevance in ECG data.
(b) A SHAP plot of statistical features calculated from ECG sequence data for the analysis ofimpact of features on the model output.
(c) A method of applying fuzzy rules to improve the interpretability of the reasoning process and resultsfor epilepsy recognition based on statistical features calculated from EEG sequence data

17



Interpretability of an FIS in disease diagnosis

17



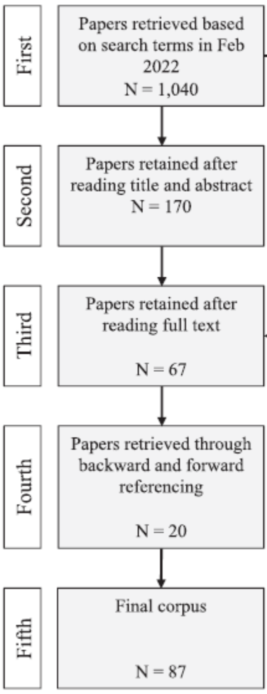
Subjective comparisons of the properties of explanation methods in disease diagnosis scenarios

- ✓ Reveals properties of algorithmiccomplexity, portability,
- ✓ And translucency remain consistent regardless of the application scenario,
- ✓ Whereas the expressive power of various explanationmethods varies across different scenarios

Literature

22

Literature search results following the PRISMA standards



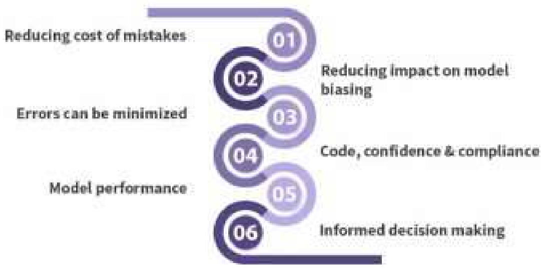
Search Terms ("Term 1" AND "Term 2")		ProQuest	Scopus	PubMed
Term 1	**Term 2**	**Results**	**Results**	**Results**
Explainable AI	Human Subjects	38	40	0
Explainable AI	Human Computer Interaction	98	459	3
Explainable AI	Human Machine interaction	23	58	1
Explainable AI	Human AI Interaction	21	148	0
Explainable AI	Human Machine Teams	6	27	0
Explainable AI	Human AI Teams	1	38	5
XAI	Human Subjects	28	38	8
Total		215	808	17
Total retained after reading title and abstract[1]		170		
Total retained after reading full text		67		

[1]Duplicates were removed after consolidation of papers across databases.

xAI

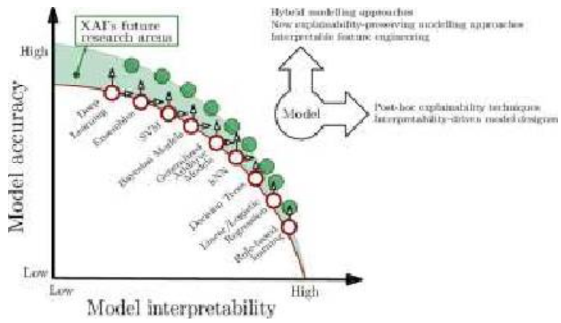
XAI benefits over AI

28



Benefits of XAI

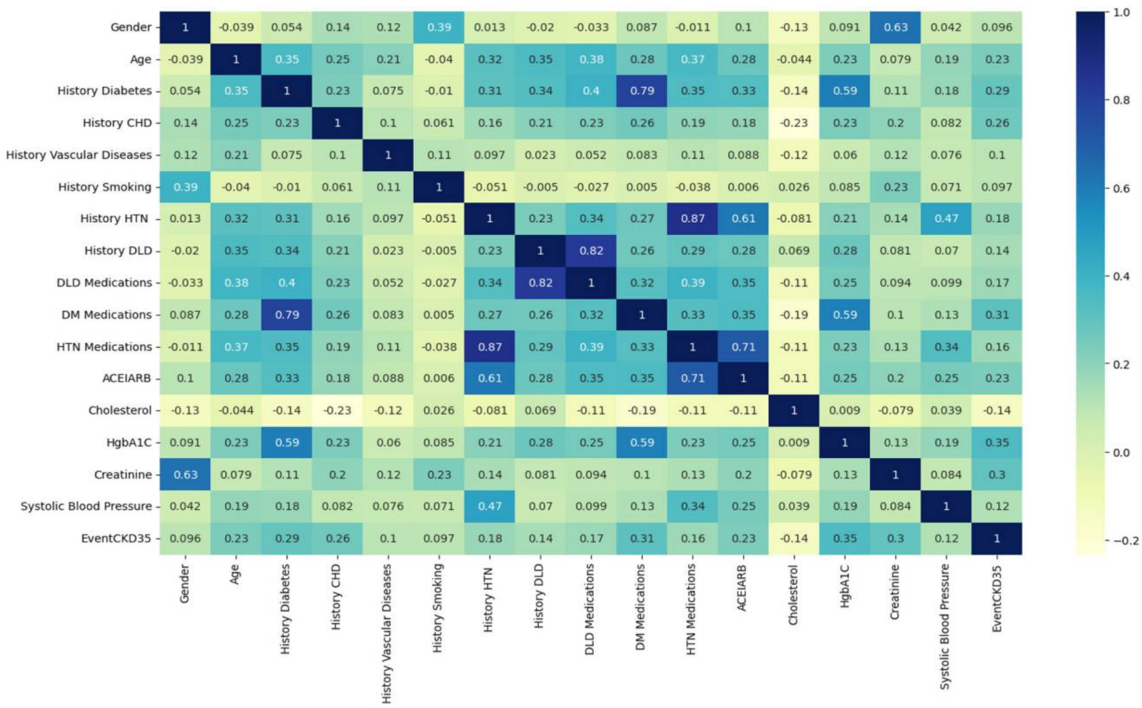
28



XAI Accuracy with
Model Interoperability

xAI Probes

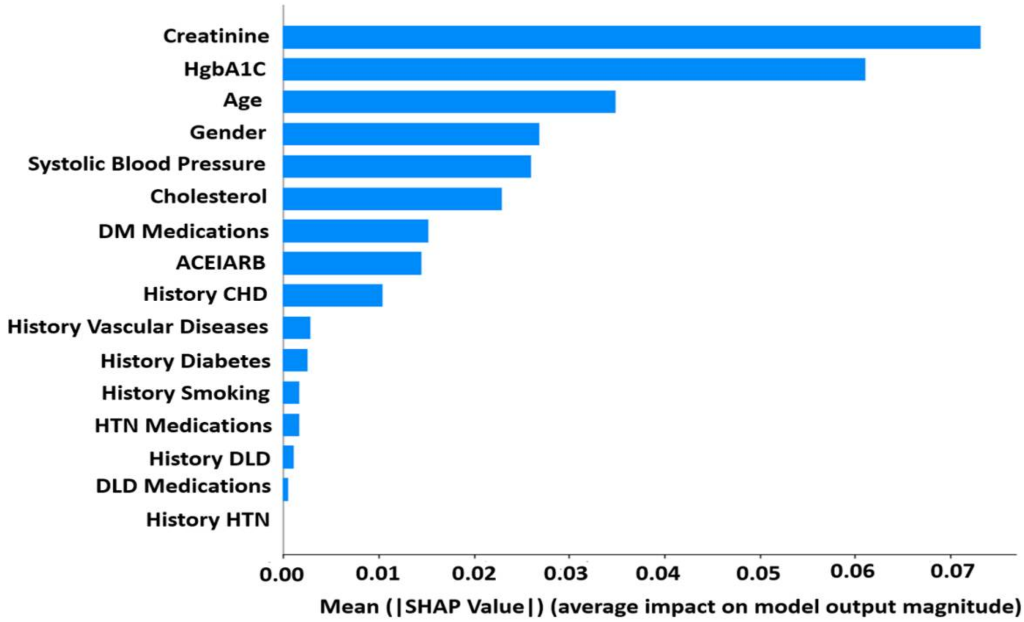
42



Heatmap of the correlation coefficient matrix.

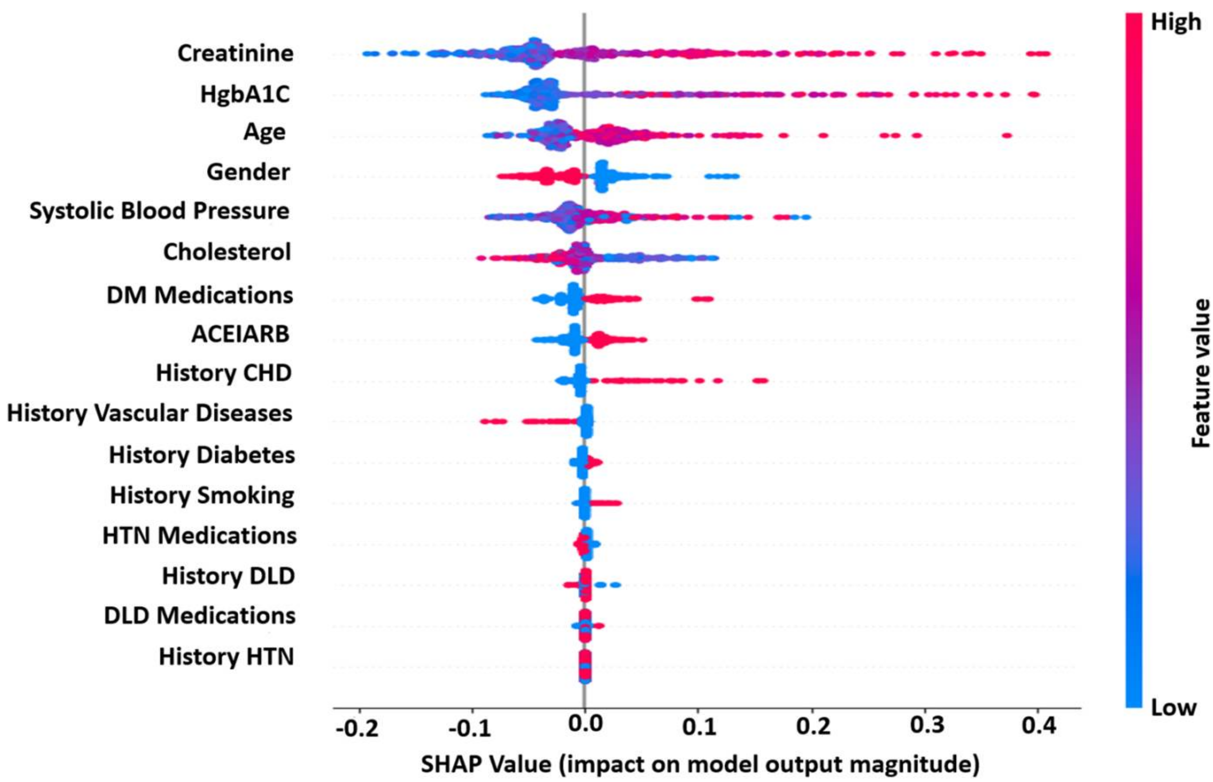
- ✓ Blue signifies a positive correlation,
- ✓ while yellow represents a negative correlation.
  - o The intensity of the color reflects the magnitude of the correlation coefficient,
  - o with more vibrant shades indicating stronger correlations.
    - Specifically, shades tending towards blue represent coefficients approaching 1,
    - while those leaning towards yellow represent coefficients approaching − 1

42



Ranking of feature importance indicated by SHAP algorithm for predicting risk of CKD

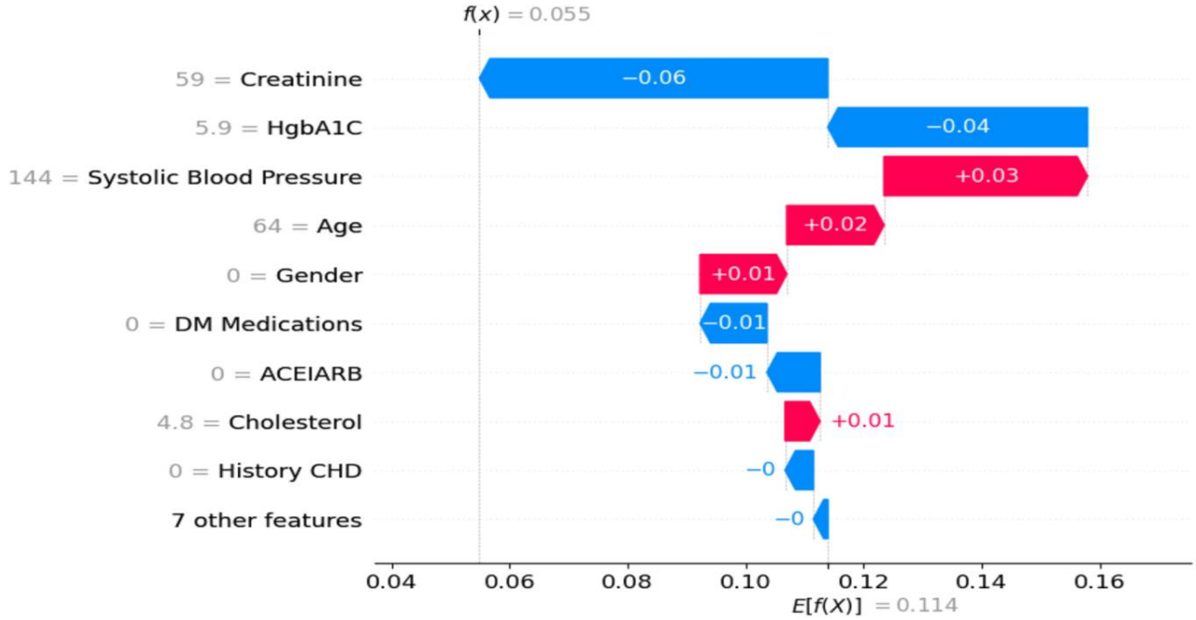
42



Beeswarm summary plot

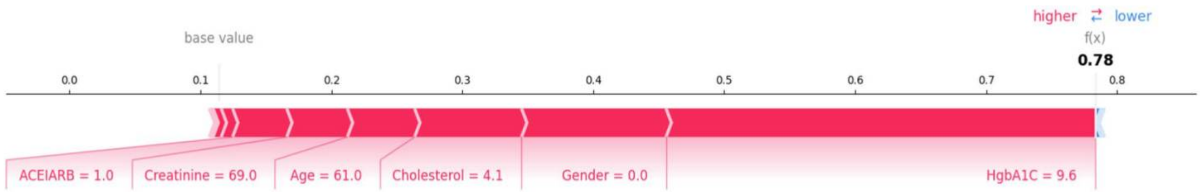
- ✓ illustrating the influence of input variables on the
  - o predictive performance ofthe XGBoost model, using SHAP values

42



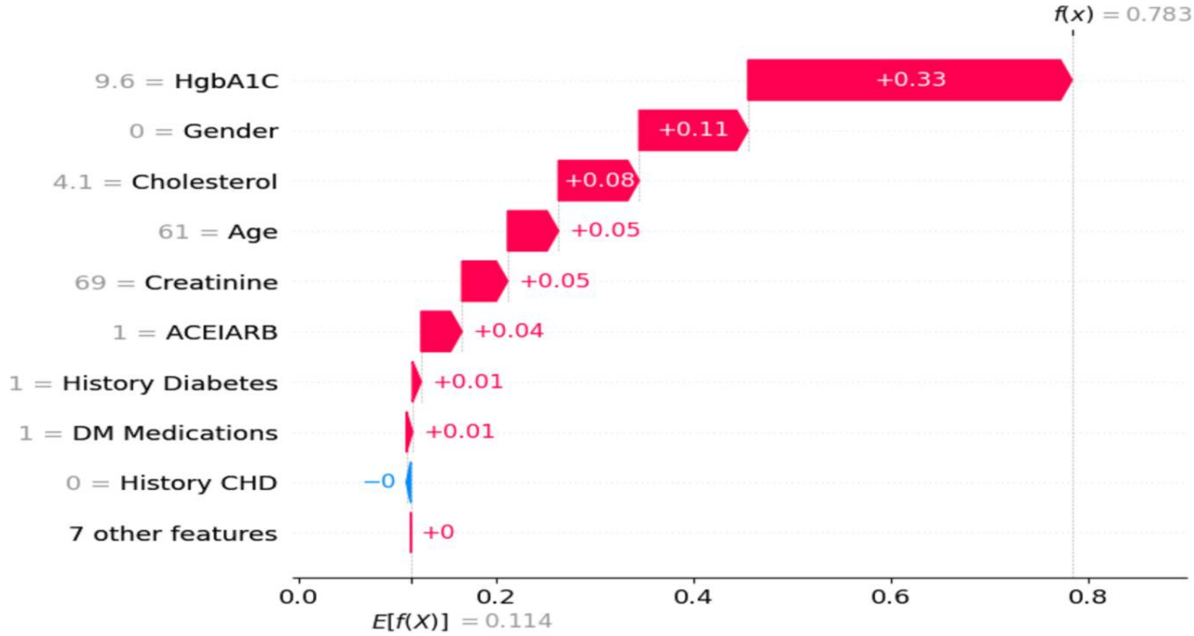
SHAP waterfall plot for patient ID 1 (true: non-CKD, predicted: non-CKD

42



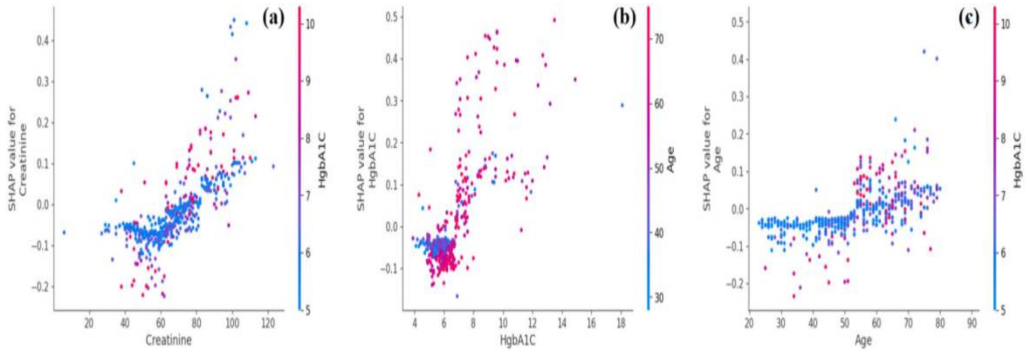
SHAP-based explainable function plot for patient ID 68 (true: CKD, predicted: CKD

42



SHAP waterfall plot for patient ID 68 (true: CKD, predicted: CKD).

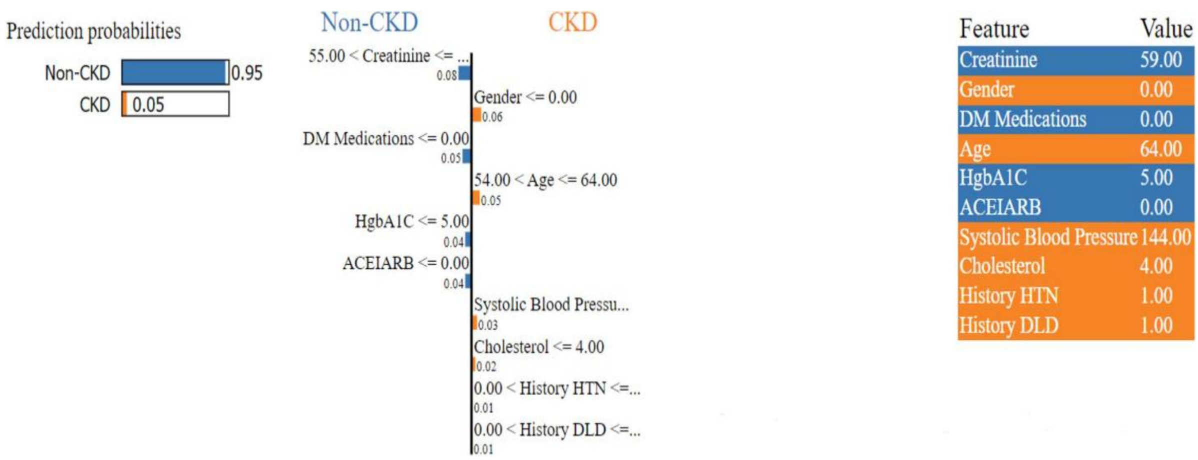
42



SHAP dependence plot for top three clinical features influencing the XGBoost model

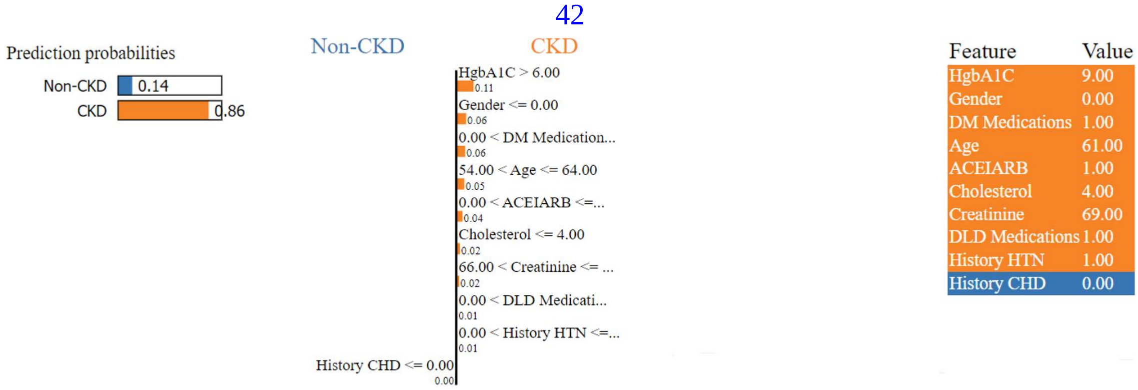
✓ (a) creatinine; (b) HgbA1C; (c) age.

42



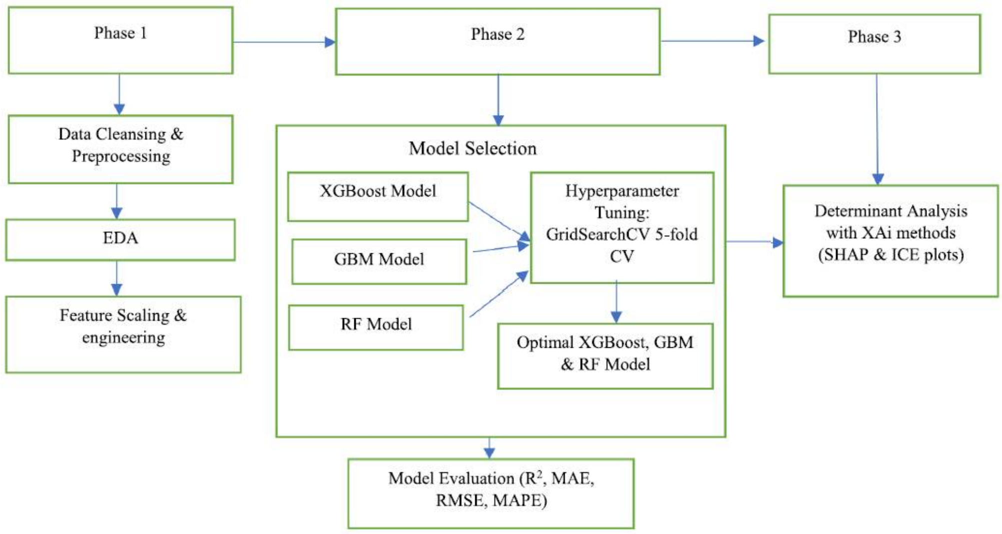
A visualization of LIME model scores for patient ID 1 using the XGBoost model.

42



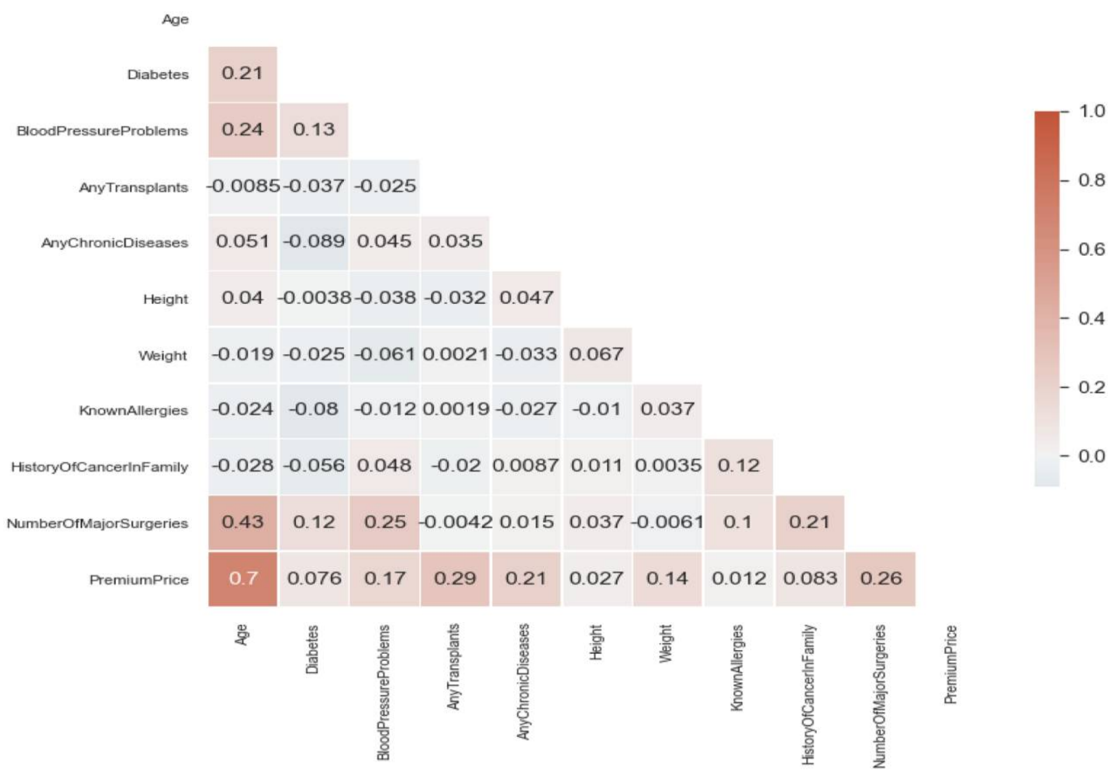
A visualization of LIME model scores for patient ID 68 using the XGBoost model

16



Workflow of the experiment

16



Correlation heatmap.

16

Advantages and disadvantages of the ML models.

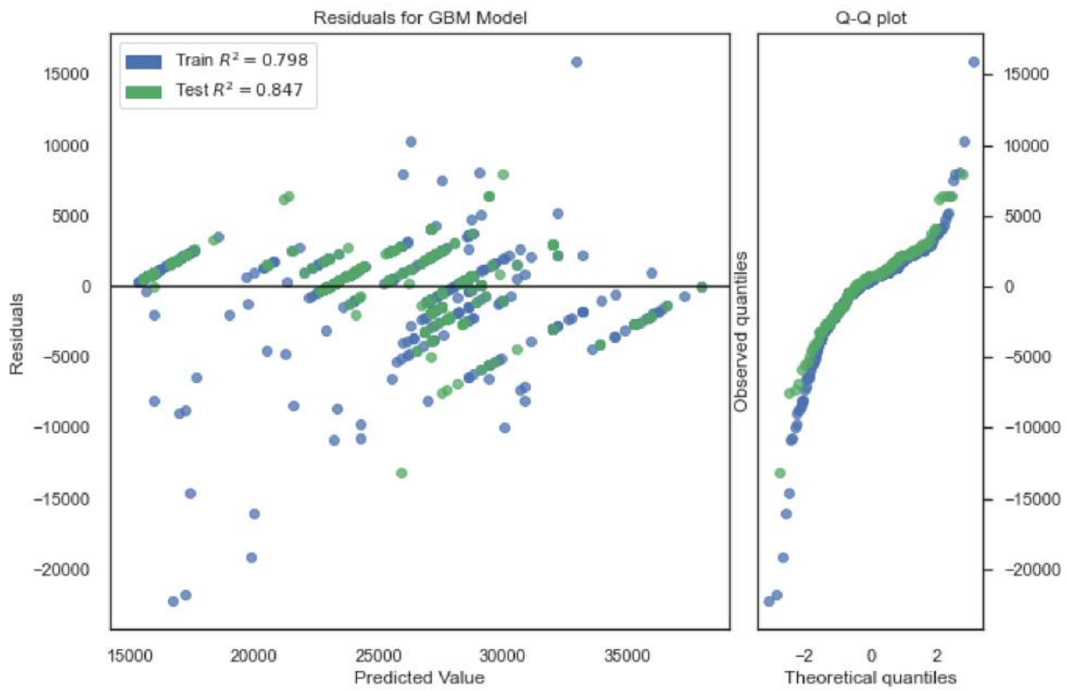
Models	Advantage	Disadvantage
RF Model	RF ensures the availability of reliable estimates for feature importance (Zhao et al., 2022).	RF models have longer computation time and consume more computational resources (Biau & Scornet, 2016).
	RF model performs well even without hyper-parameter tuning (Gomes et al., 2017).	Prone to overfit with noisy data (Hoarau et al., 2023).
	The presence of missing values does not hinder RF (Tyralis et al., 2019).	RF is relatively hard to interpret (Marchese Robinson et al., 2017).
XGBoost Model	XGBoost performs well with little or no feature engineering and can handle missing data (Kang et al., 2020).	If not properly tuned, XGBoost is more likely to overfit (Priscilla & Prabha, 2020).
	XGBoost is renowned for its computational speed, model performance, and is well-known to handle large-sized datasets (Chen, et al., 2015).	It is harder to tune as there are too many hyper-parameters (Zivkovic et al., 2022).
GBM Model	Improved convergence speed without a significant decrease in accuracy (Feng, Xu, & Tao, 2018)	Achieving a balance between performance and generality has posed a challenge for GBMs (Luo, Wei, Man, & Xu, 2022).
	Gradient boosting of regression trees produces competitive, highly robust, interpretable procedures for both regression and classification, especially appropriate for mining less than clean data (Friedman, 2001).	Like in XGBoost, GBM has many hyperparameters that need proper tuning (Anghel et al., 2018; Kiatkarun & Phunchongharn, 2020).

16

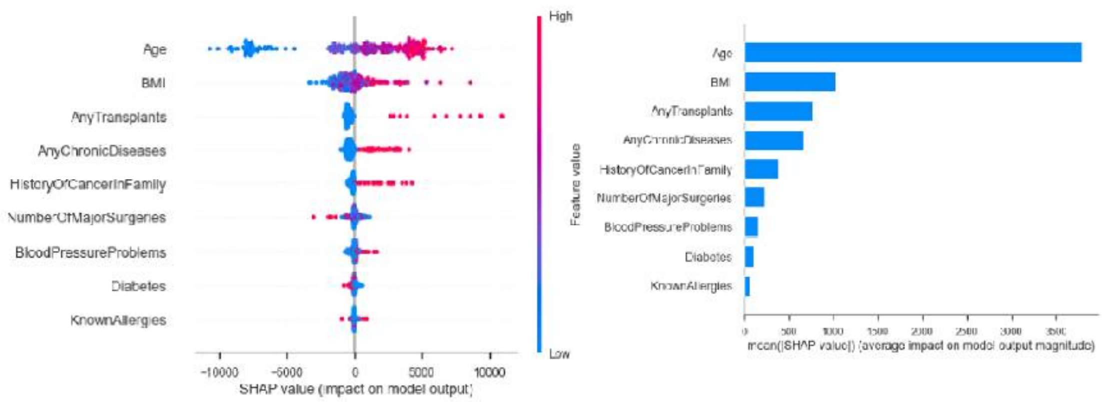
Advantages and disadvantages of SHAP.

XAI Method	Advantage	Disadvantage
SHAP Analysis	Global interpretability— SHAP helps determine whether each variable is positively or negatively related to the target variable (Lundberg & Lee, 2017).	SHAP comes with computational complexity and consumes huge computation resources (Lin & Gao, 2022).
	Local interpretability— all features are represented with a SHAP value (Stiglic et al., 2020).	SHAPley values cause extrapolation to low-density areas for dependent features (Lundberg & Lee, 2017).
	SHAP calculates the contribution of each feature to the prediction (Teoh, et al., 2022).	Regardless of how small the change may be, every feature that changes the prediction is attributed a SHAPley value other than zero (Janizek et al., 2018).
ICE Plots	A fitted model's ICE plot can reveal heterogeneous relationships between predictors and predicted values by visualizing the map between predictors and predicted values (Casalicchio et al, 2019).	According to the joint feature distribution, some points on the ICE curves might be invalid data points if the feature of interest is correlated with the other features (Molnar et al. 2022).
	The process of creating an ICE plot is extremely straightforward (Goldstein et al. 2015).	Generating ICE plots can be time-consuming, especially for large datasets or complex models (Molnar, 2020).

16



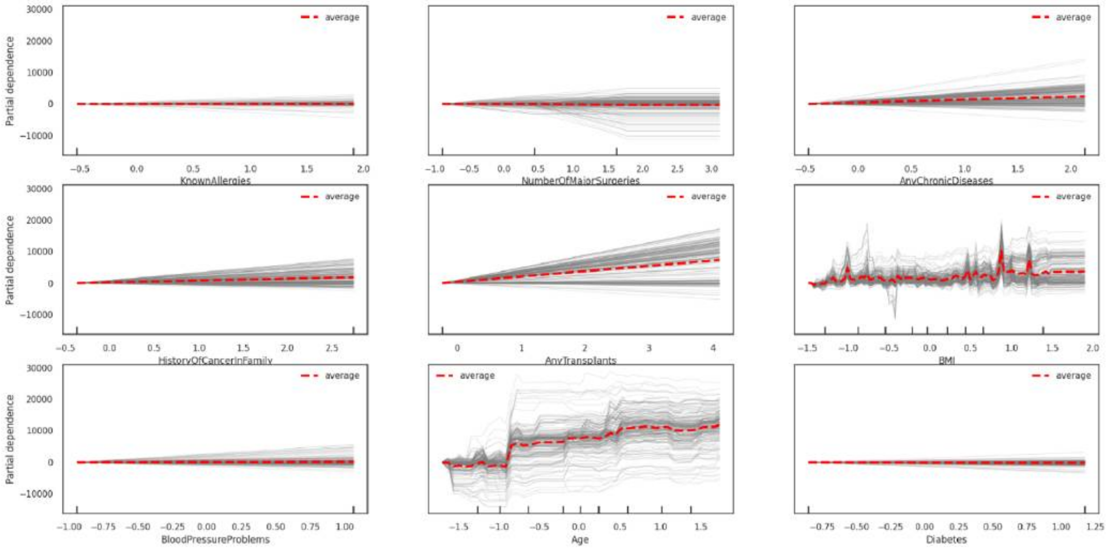
16-



SHAP summary plot

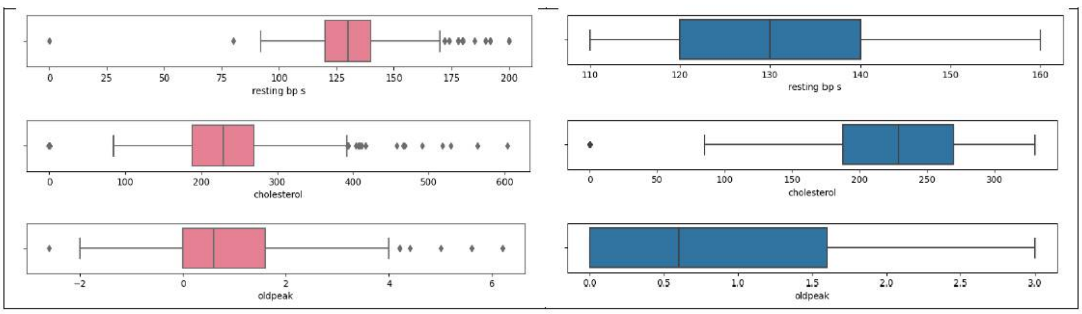
- ✓ Relative importance of each feature in the XGBoost model:
- ✓ (a) SHAP summary plot; (b) relative importance

16



C-ICE (Central-Individual Conditional Expectation) plot
for XGBoost model

08



Distribution of features with outlier and after replacing them

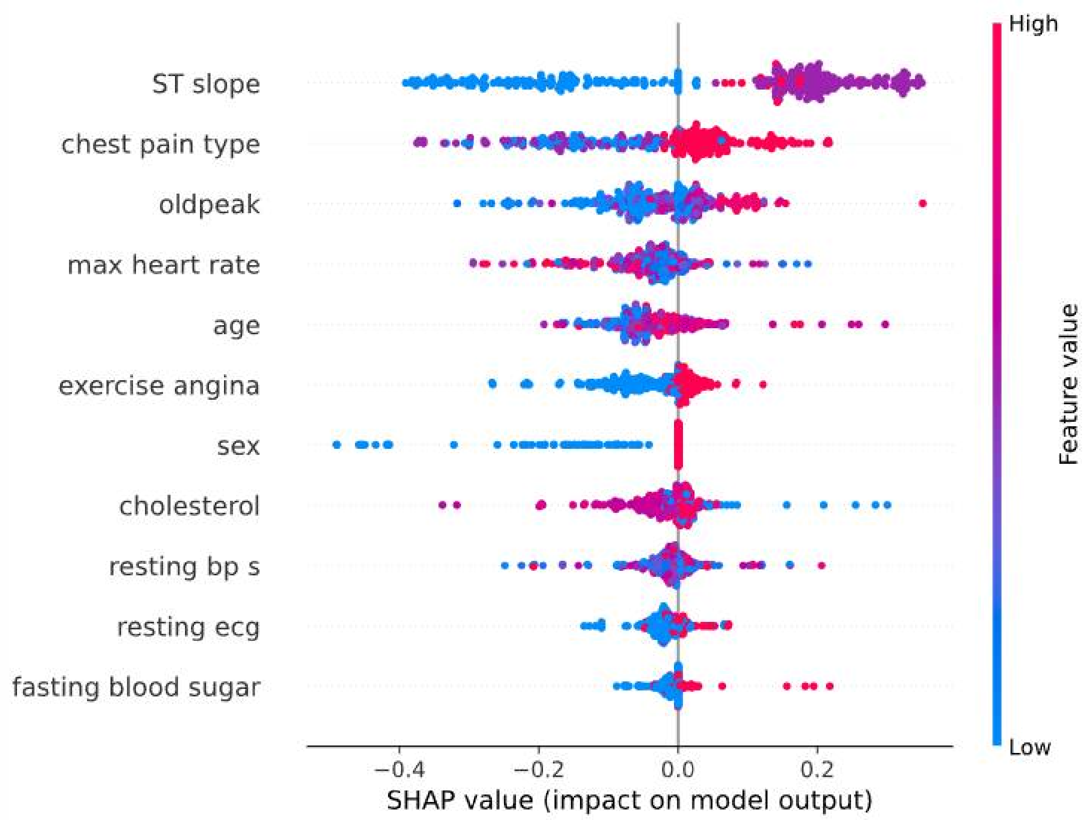
08

	Accuracy	Precision	Recall	F1
LR	85	84	84	84
DTC	88	87	88	87
NB	85	85	85	85
XGBoost	92	92	92	92
RFC	93	93	93	93

ANN	90	90	90	90
KNN	88	88	88	88
SVM	89	89	89	89
DTC + LR + NB	89	89	89	89
LR + XGB + KNN	92	92	92	92
RFC + XGB + SVC	**94**	**94**	**94**	**94**

Performance Comparison of Models

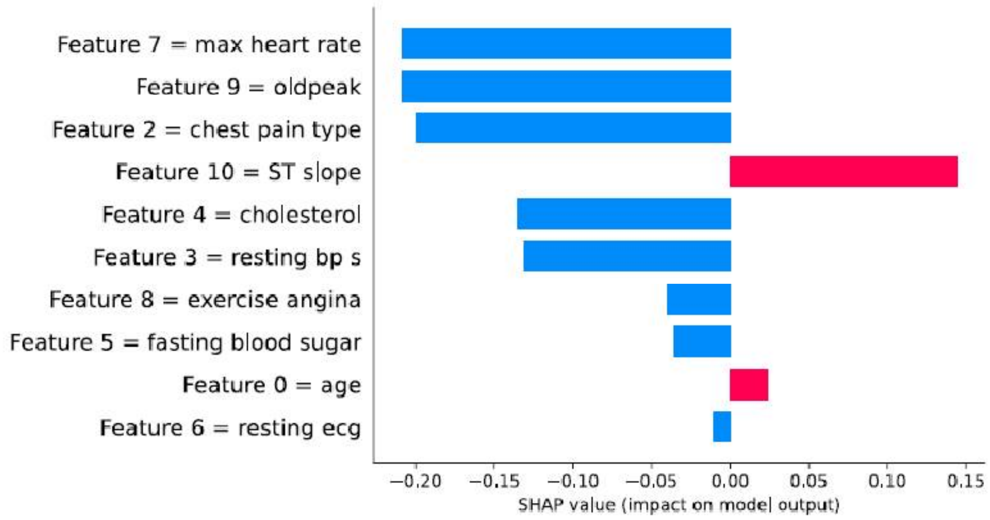
08



SHAP Summary
Plot for Best Ensemble Model

08

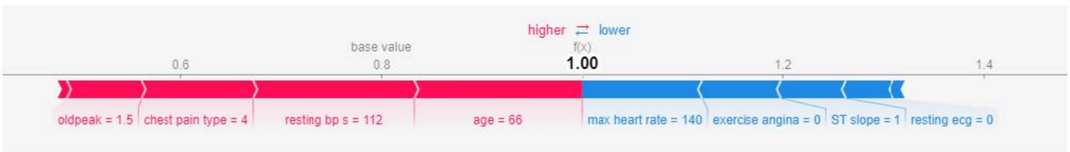
(a) Force plot



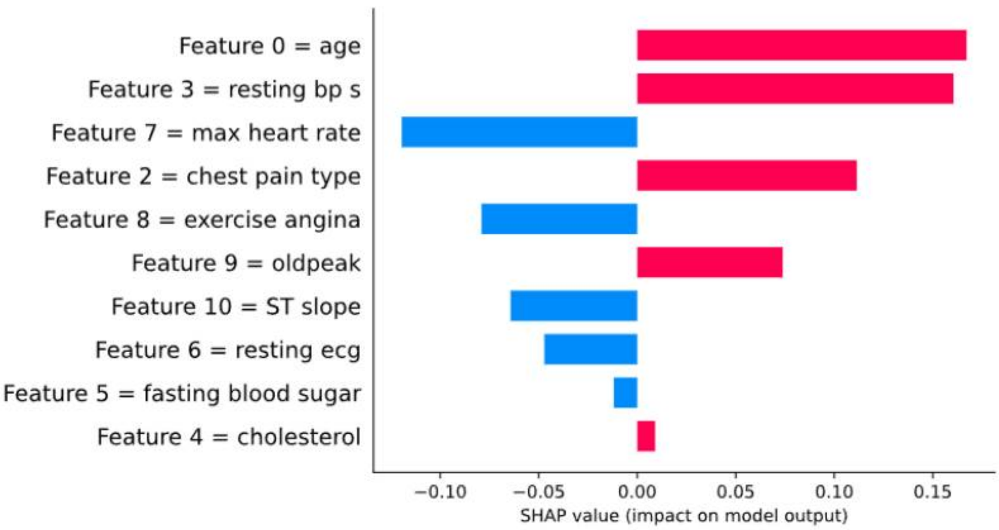
(b) Bar plot

SHAP Values in class 0 prediction

08



(a) Force plot

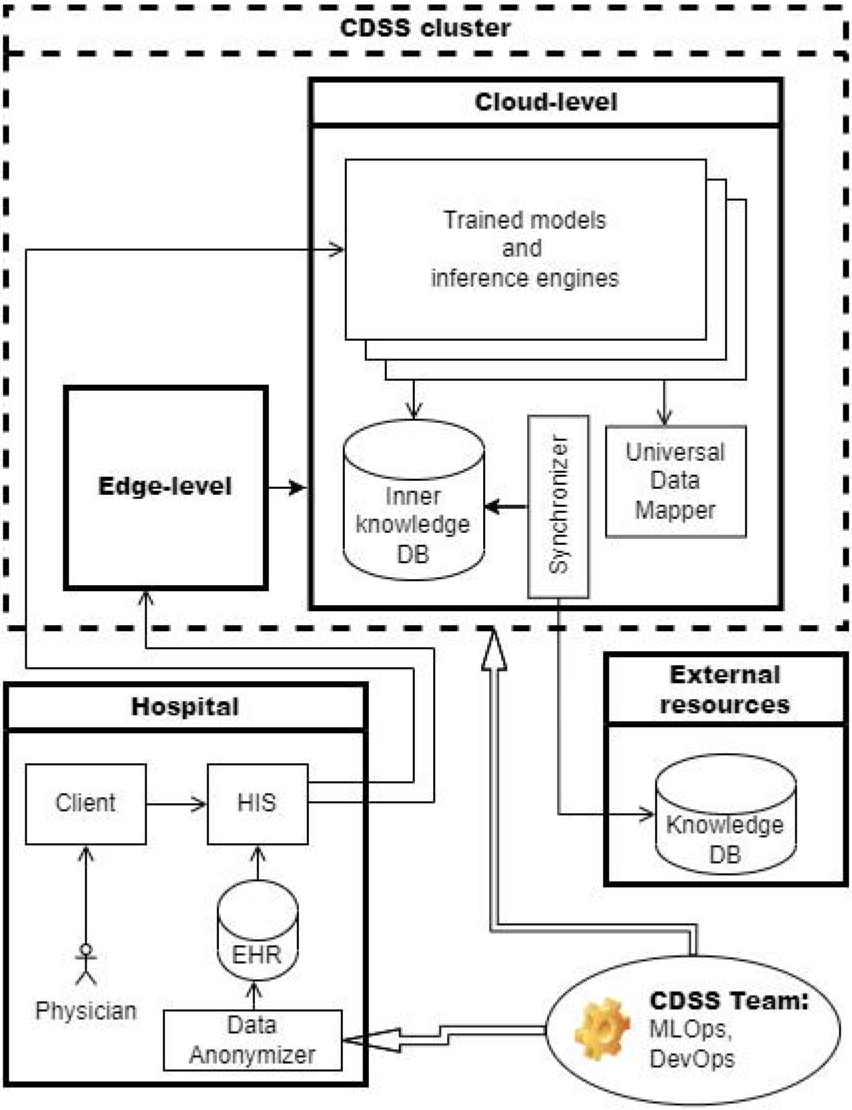


(b) Bar Plot

SHAP Values in class 1 prediction

Clinical
Decision Support System (DSS)

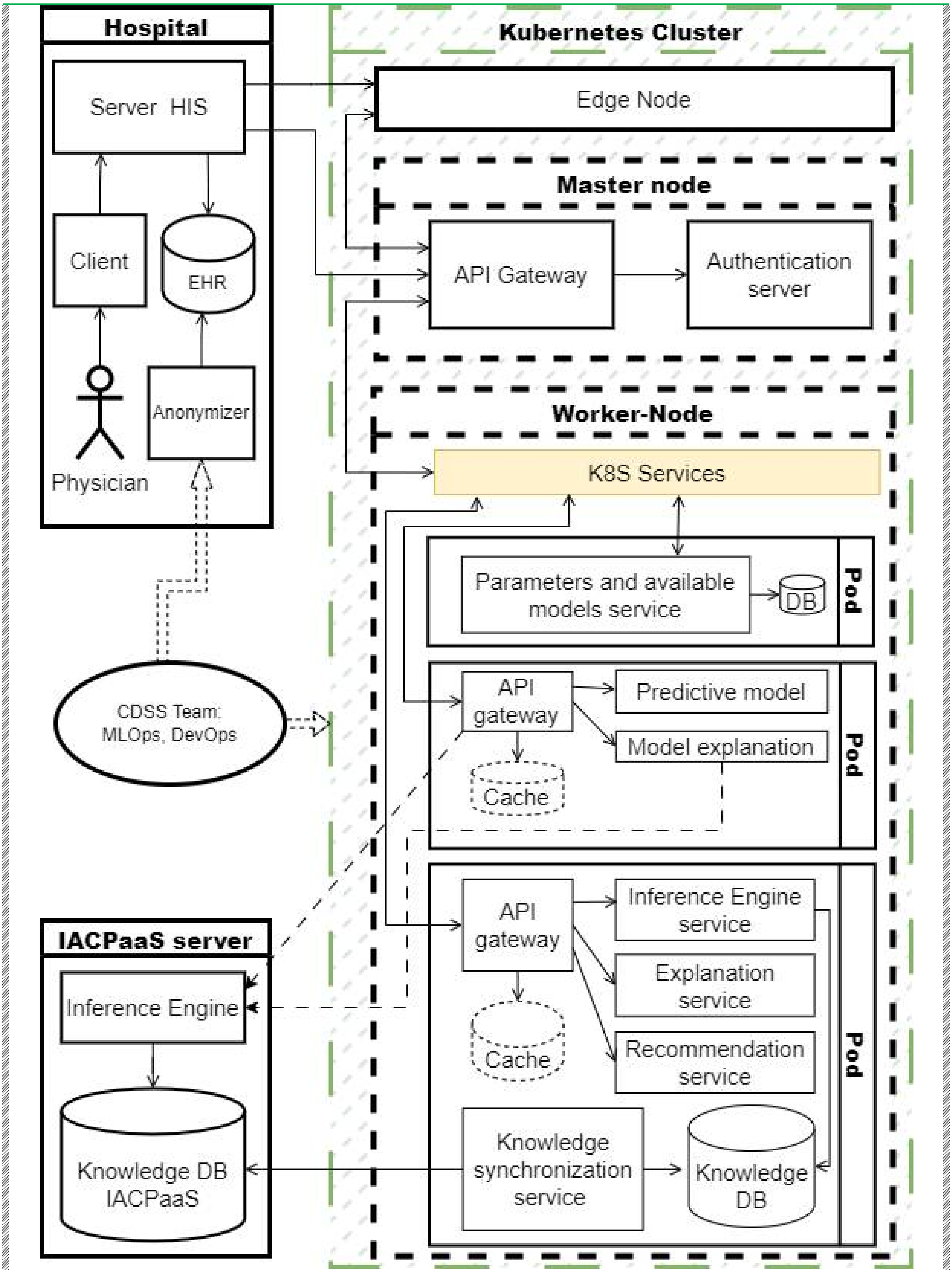
05



Clinical decision support system

✓ high-level architecture

05

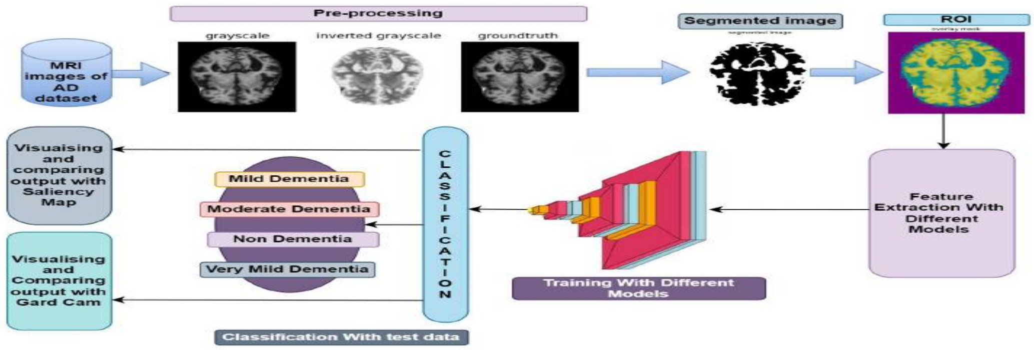


Clinical decision support system

✓ low-level architecture

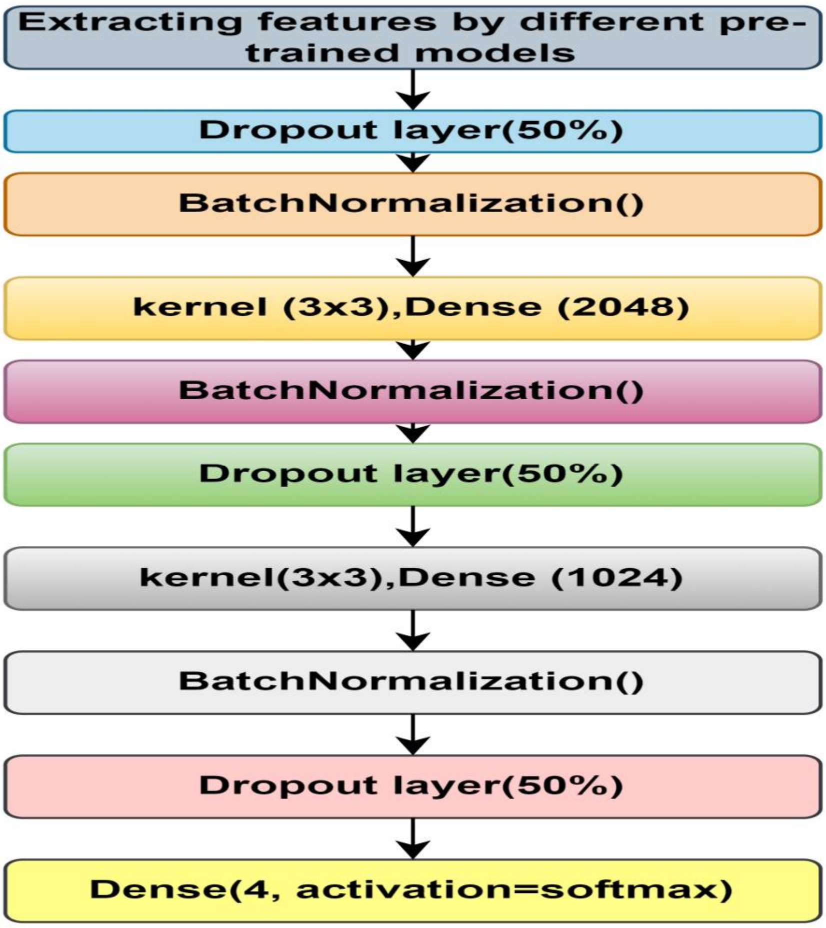
Alzheimer's disease

02



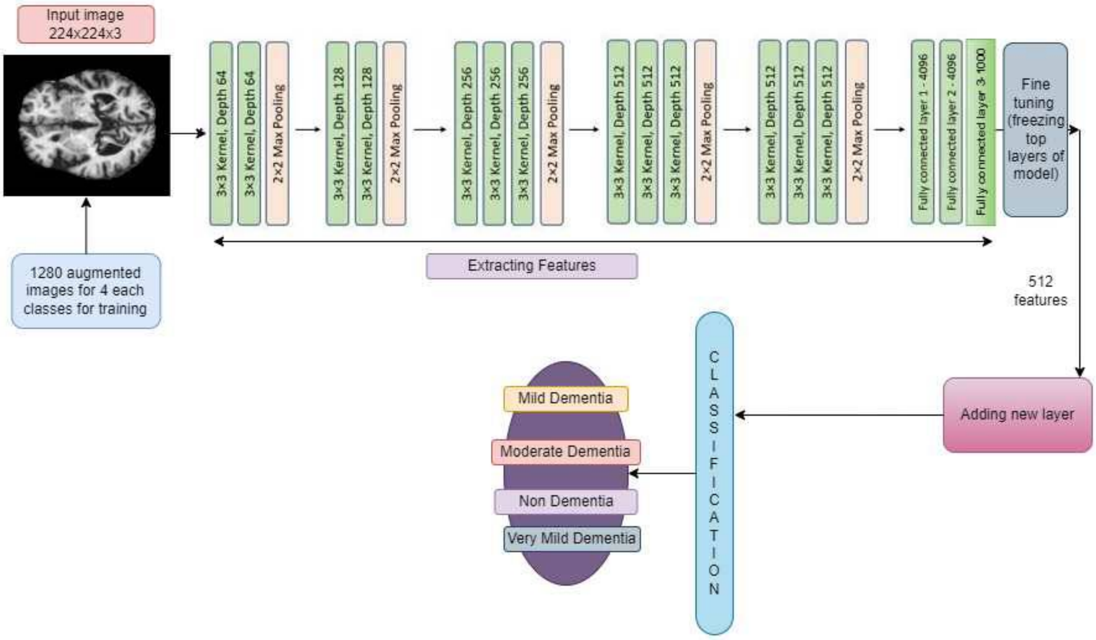
Workflow diagram for diagnosing MRI images for the detection of AD.

02



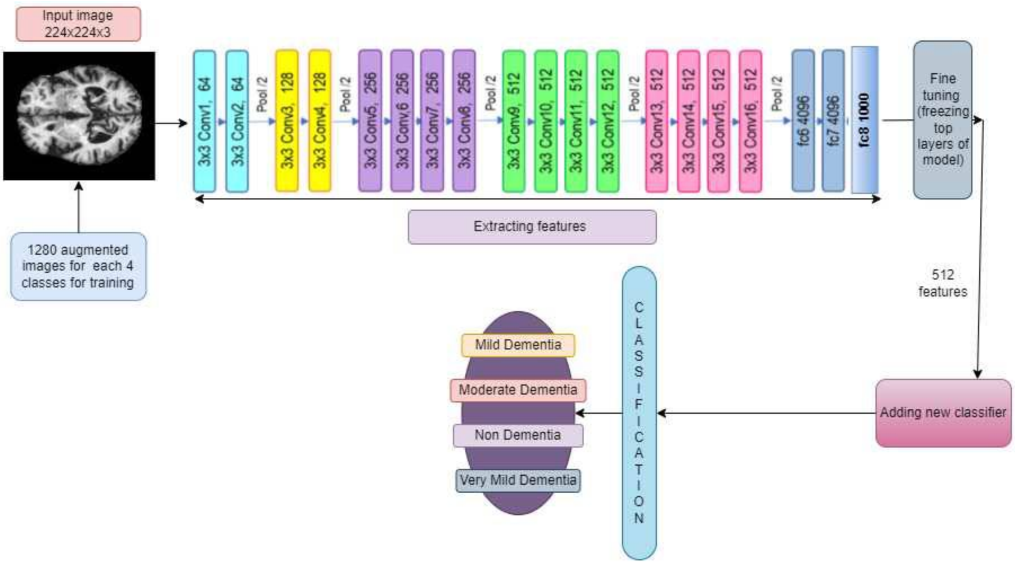
Addition of new layers after feature extraction

02



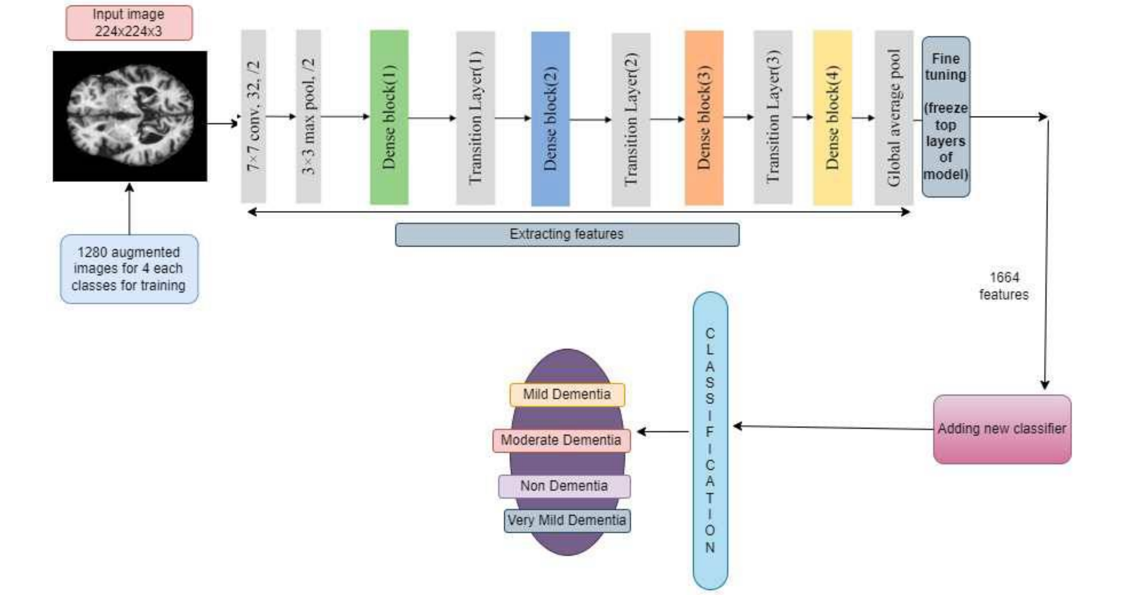
Addition of new layers after feature extraction using VGG16

02



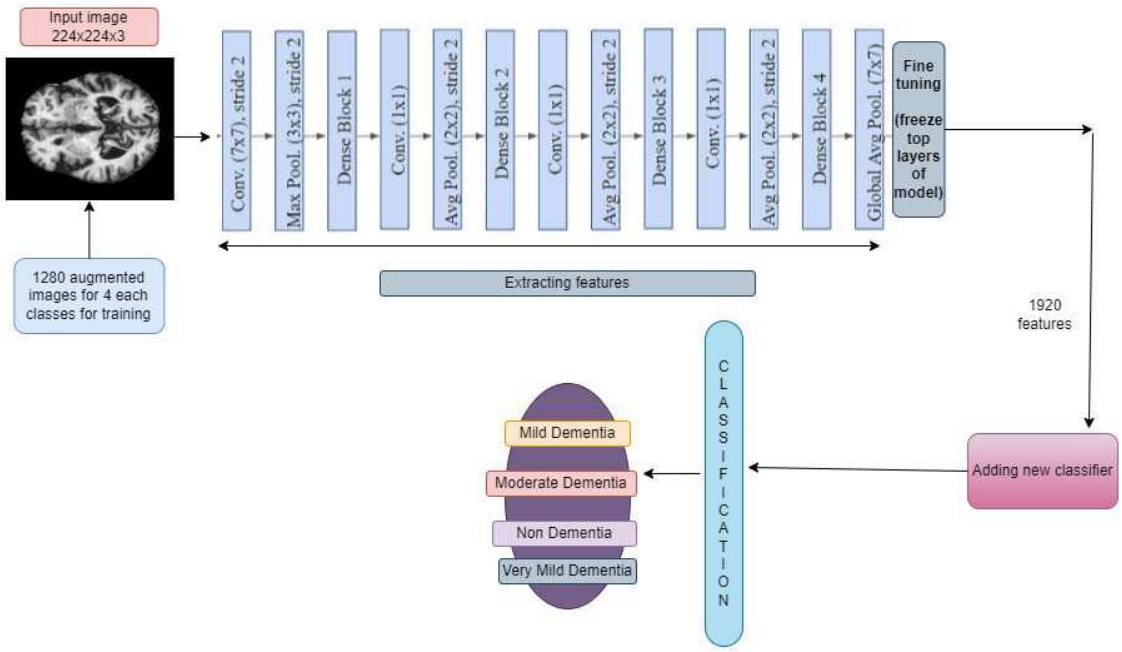
Addition of new layers after feature extraction using VGG19

02



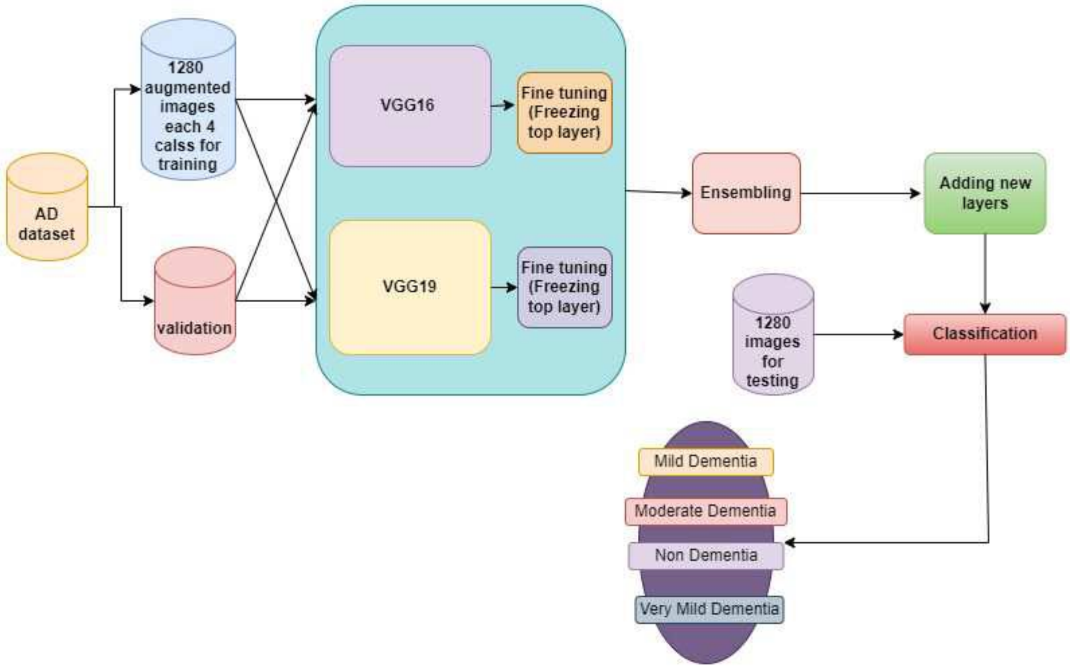
Addition of new layers after feature extraction using DenseNet169

02



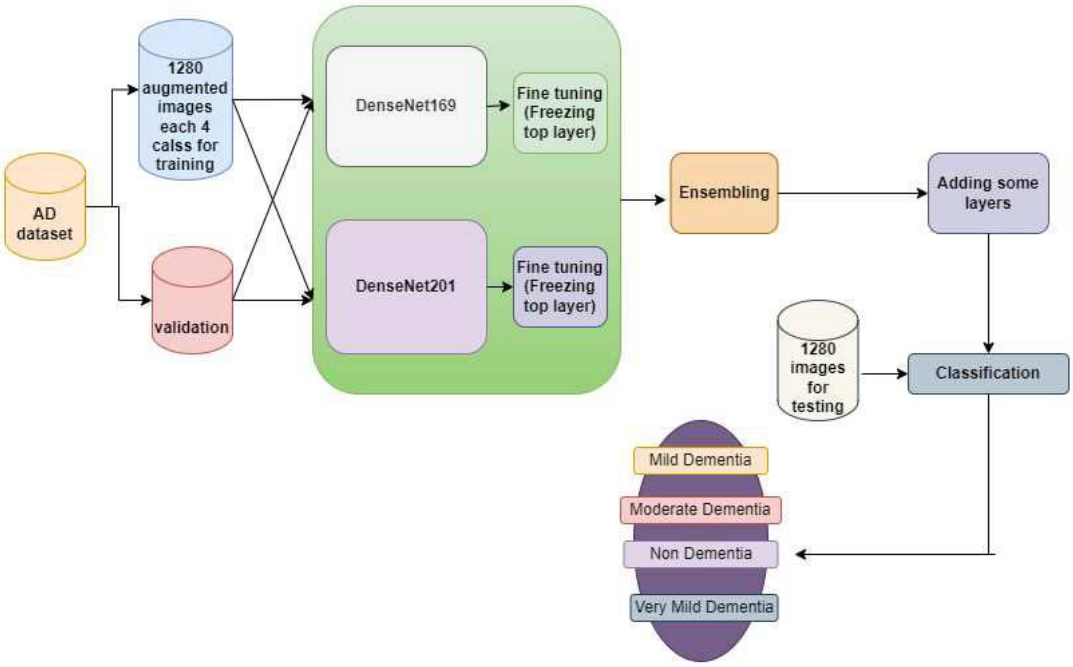
Addition of new layers after feature extraction using DenseNet201

02



Ensemble-1 model architecture

02

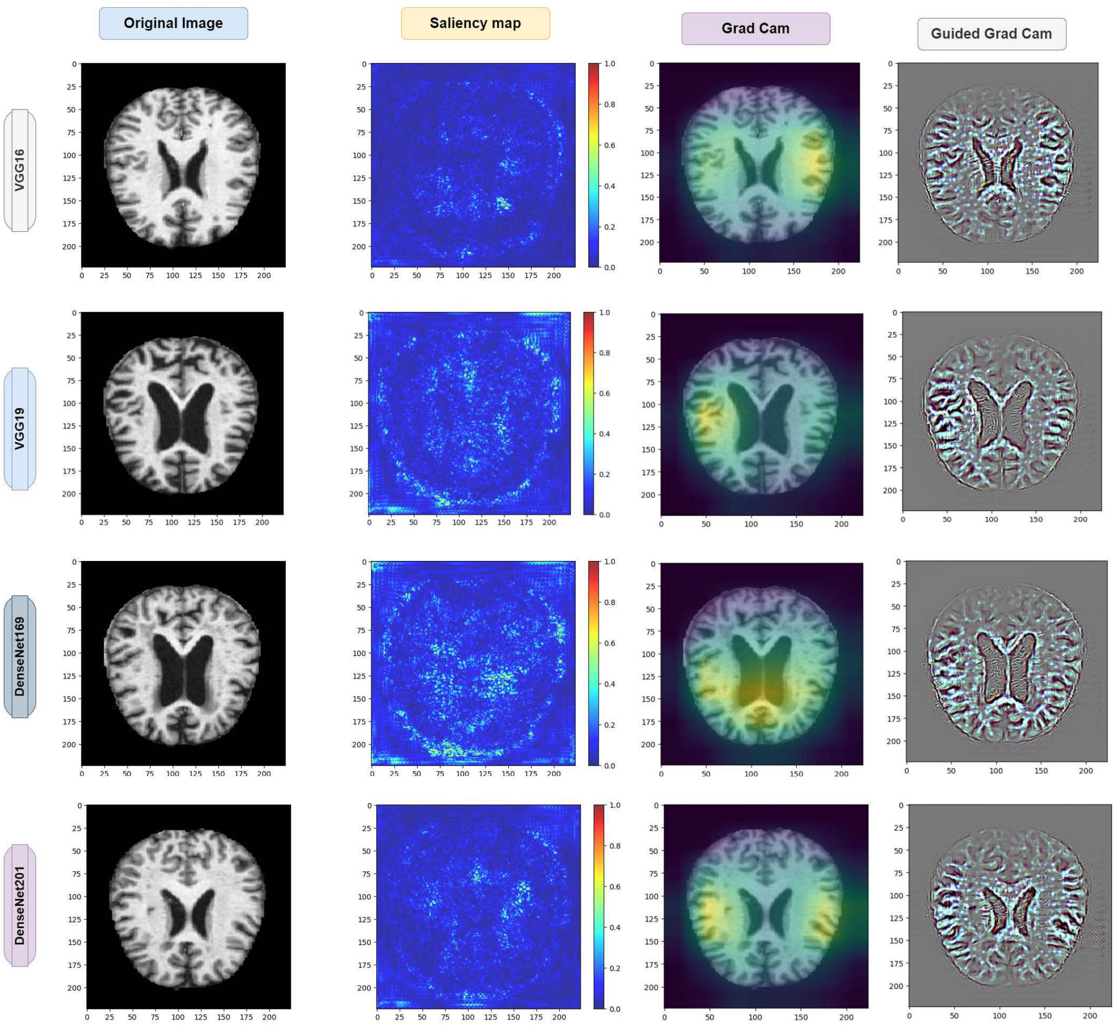


Ensemble-2 model architecture

Description of the proposed model architecture

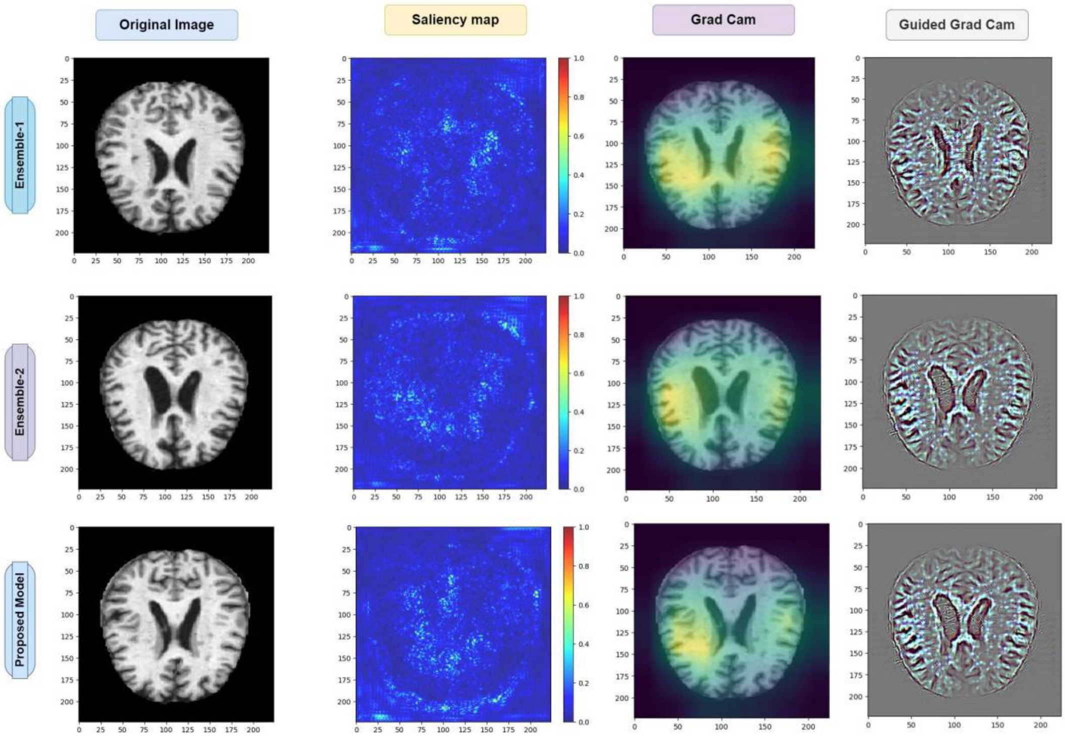
Model Content	Details
Input Image Size	224 × 224 × 3, with 5120 training images and 1280 images in each class
Feature extraction	Using EfficientNet with 1280 features
First Convolution Layer	32 filters; size = 3 × 3; ReLu; Padding = 'Same'
First Max Pooling Layer	Pooling Size: 2 × 2
Second Convolution Layer	64 filters; size = 3 × 3; ReLu; Padding = 'Same'
Second Max Pooling Layer	Pooling size: 2 × 2
Third Convolution Layer	128 filters; size = 3 × 3; ReLu; Padding = 'Same'
Third Max Pooling Layer	Pooling size: 2 × 2
Fourth Convolution Layer	256 filters; size = 3 × 3; ReLu; Padding = 'Same'
Fourth Max Pooling Layer	Pooling Size: 2 × 2
Fifth Convolution Layer	512 filters; size = 3 × 3; ReLu; Padding = 'Same'
Fifth Max Pooling Layer	Pooling Size: 2 × 2
Fully Connected Layer	4096 nodes; ReLU
Dropout Layer	50% Neurons dropped randomly
Dense_1 Layer	8320 nodes; ReLu
Dense_2 Layer	516 nodes; ReLu
Output Layer	Four nodes; Softmax activation
Optimization Function	Adam optimization
Learning Rate	0.001
Loss Function	Categorical cross entropy

02



Visual depiction of the saliency map and grad-CAM results for the pretrained model

02



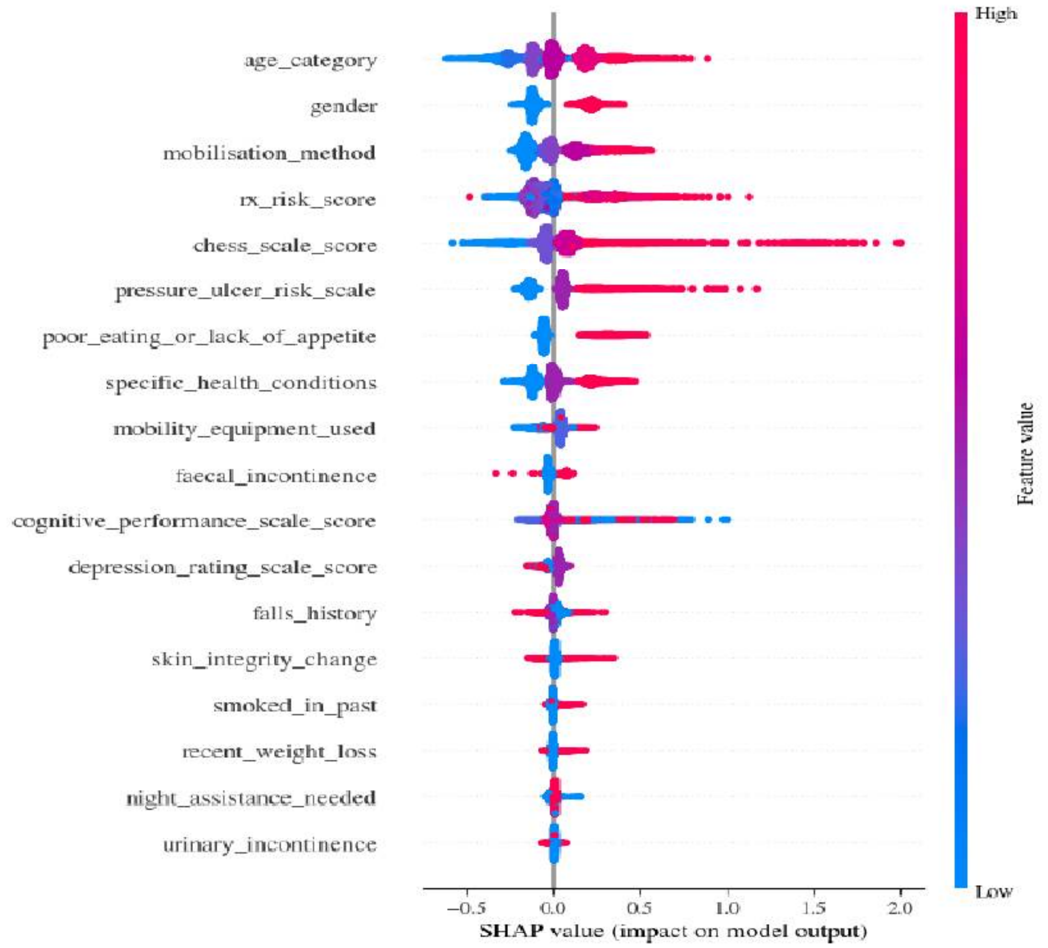
Visual depiction of the saliency map and grad-CAM results for the ensemble and proposed model.

02

Comparison of the proposed method with the state-of-the-art methods

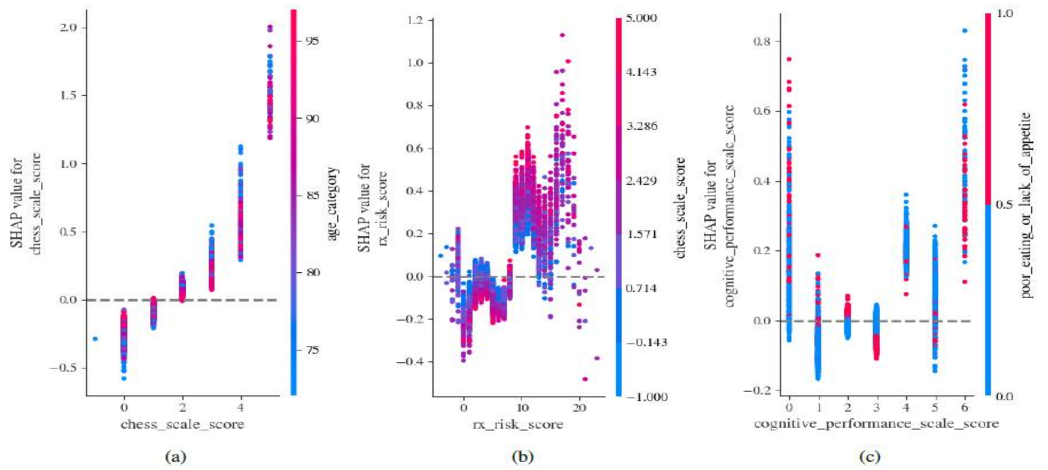
Paper	Classifier	Best Score (Accuracy)	XAI Method	Dataset
[15]	Support Vector Machines, KNN, MLP	91.4%	LIME, SHAP	Dementia dataset
[16]	CNN	95.4%	HAM, PCR	MRI scans ADNI
[17]	Graph Neural Network (GNN)	53.5 ± 4.5%	GNN Explainer	ADNI
[18]	EfficientNetB0	80%	Occlusion Sensitivity Mapping	MRI scans OASIS
[19]	3D CNN	-	Saliency Map, LRP	18F-FDG PET
[20]	KNN, RF, AdaBoost, Gradient Boosting Bernouli NB, SVM	91%	DT	Cognitive and and PET images
[21]	3D CNN	76.6%	3D Ultrametric Contour Map, 3D Class Activation Map, 3D GradCAM	ADNI
[22]	3D CNN	77%	Sensitivity Analysis Occlusion	MRI, PET
Proposed Method	VGG16, VGG19, DenseNet169, DenseNet201, Ensemble 1 (VGG16, VGG19) Ensemble 2 (DenseNet169, DenseNet20), Proposed model (EfficientNetB3 & CNN)	96%	Saliency maps, Grad-CAM (Gradient-weighted Class Activation Mapping)	MRI scans OASIS

01



Predictor importance summary plot for the XGBoost model

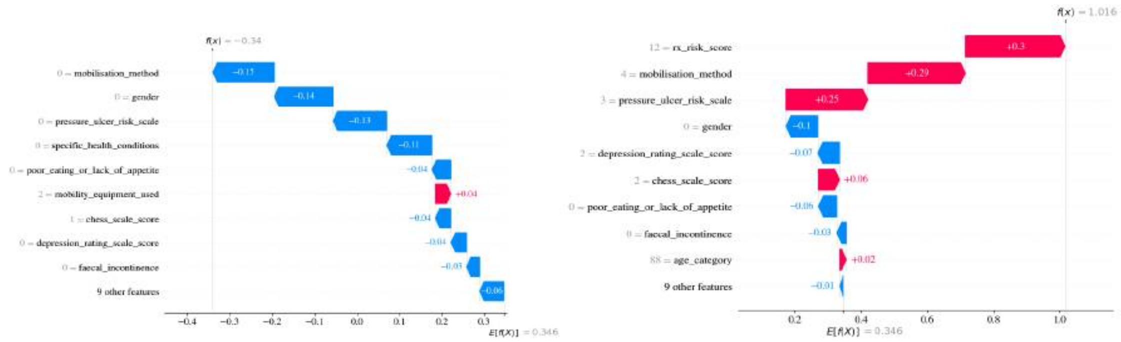
01



Dependence plots for XGBoost model

- ✓ Show pairwise interactions between a selection of predictors and
- ✓ How their interactions affect patient risk predictions

1

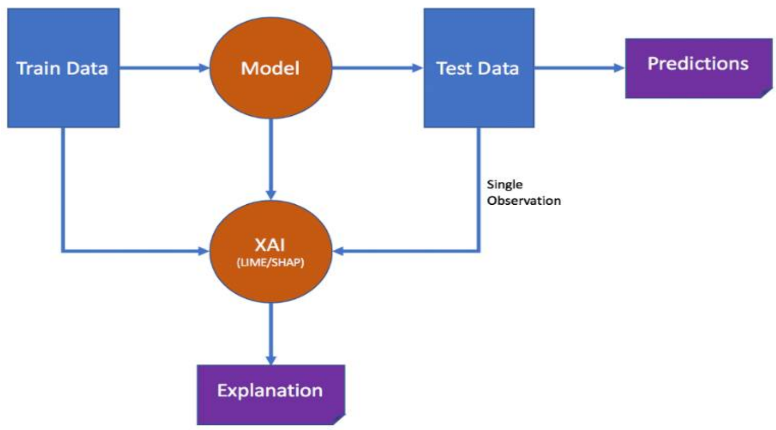


(a) Patient A (b) Patient B

Waterfall plots from the XGBoost model

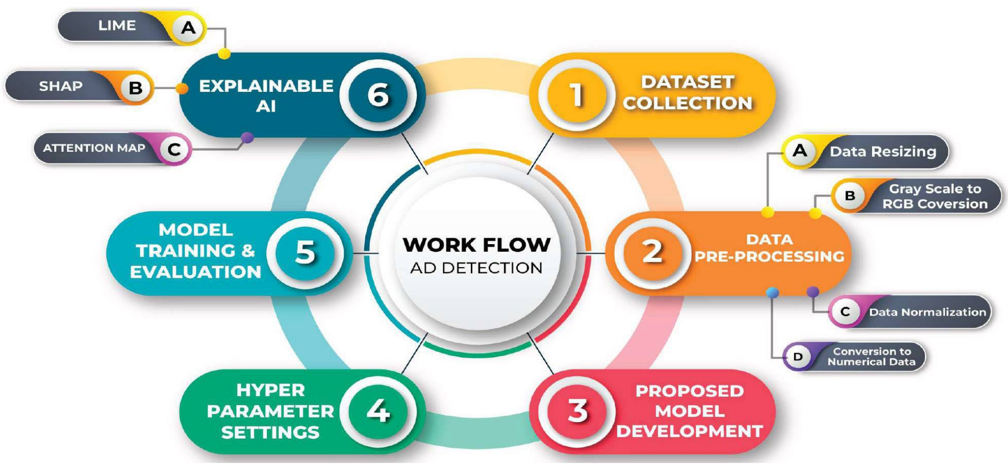
! Show interpretability for patient-level predictions using two hypothetical examples

03



Overview of the method

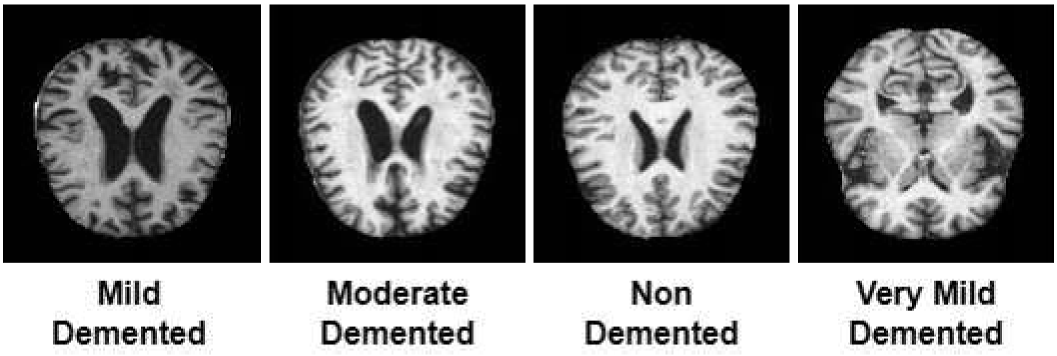
04



Workflow of the proposed methodology

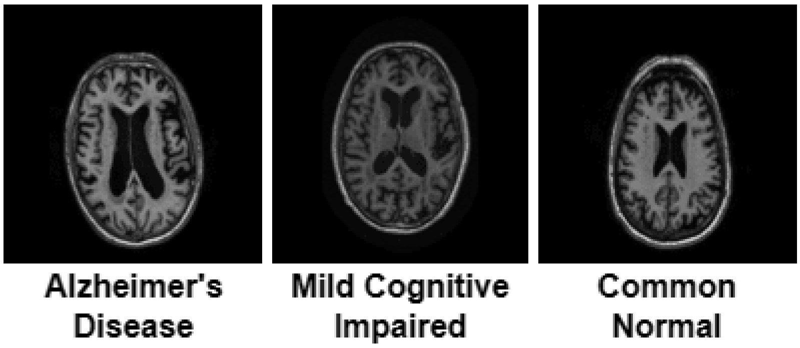
✓ Show-casing the sequential steps involved in the research process

04



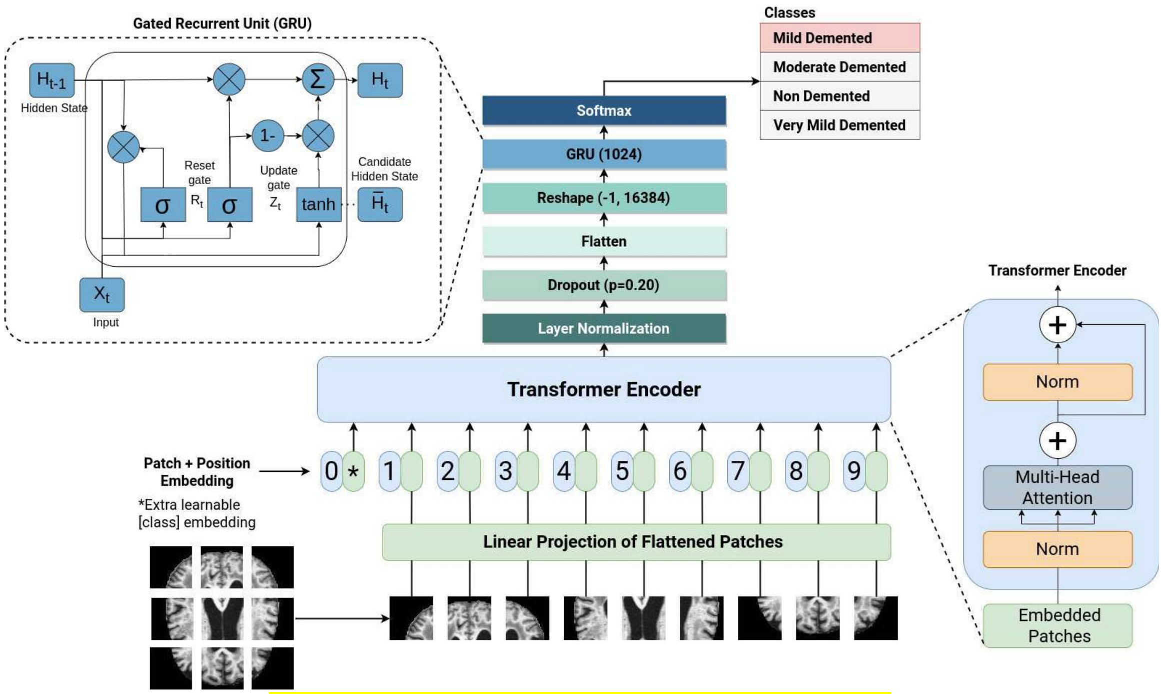
Sample images of the dataset 1.

04



Sample images of the dataset 3.

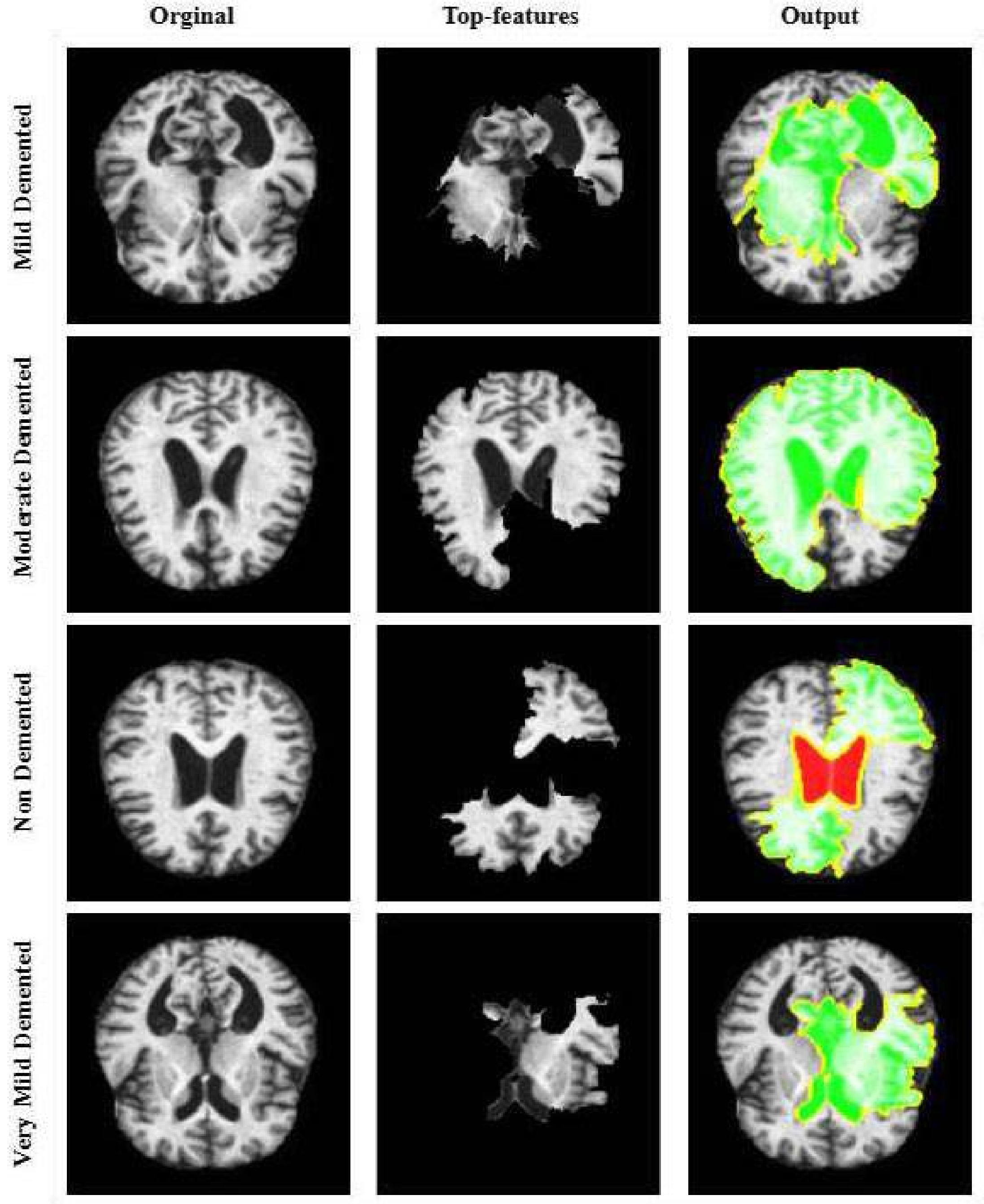
04



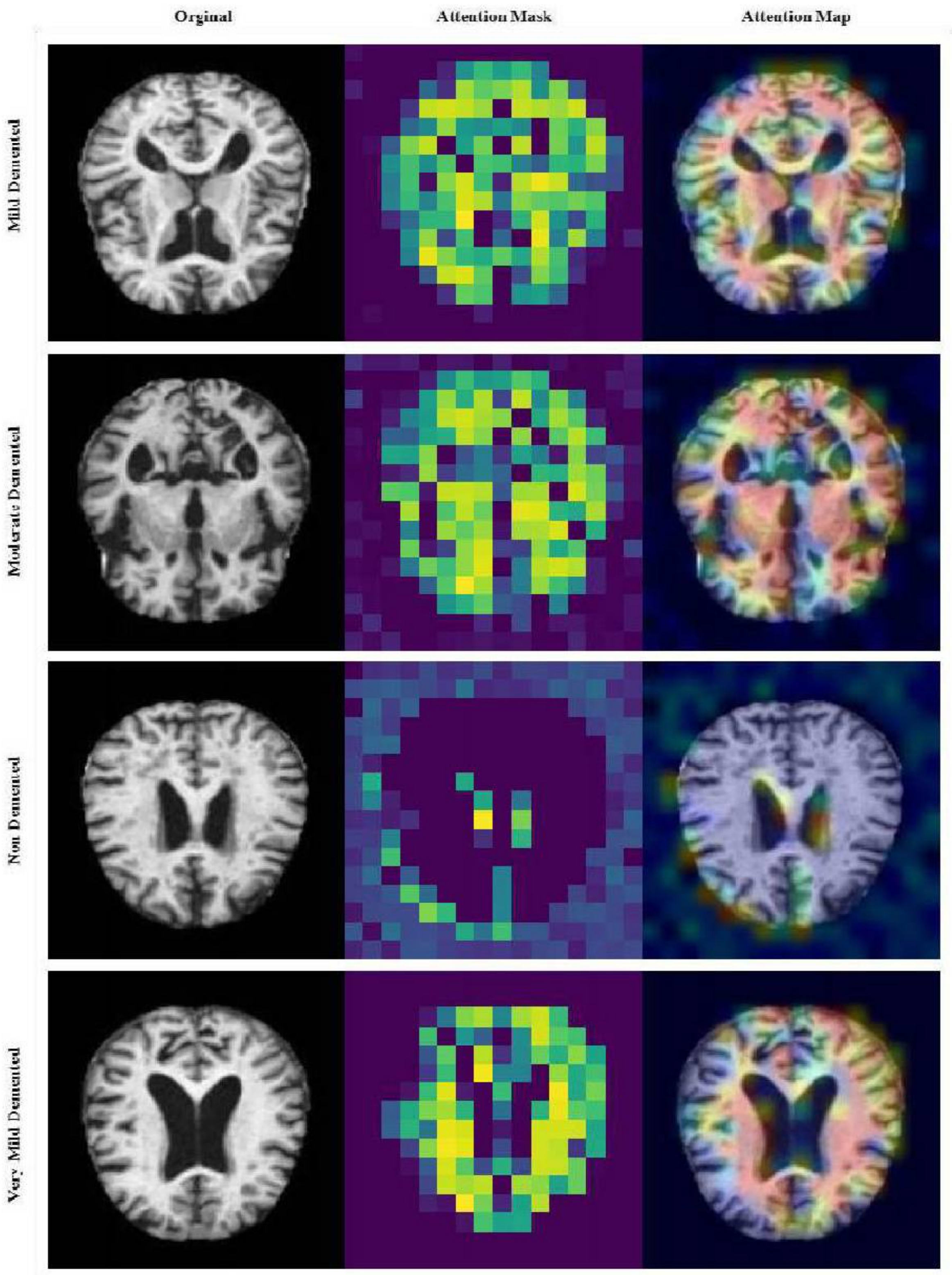
Architecture of the hybrid ViT-GRU model illustrate

Interconnected components and their functional relationship in the proposed model

04



LIME for dataset 1 (Top three features)

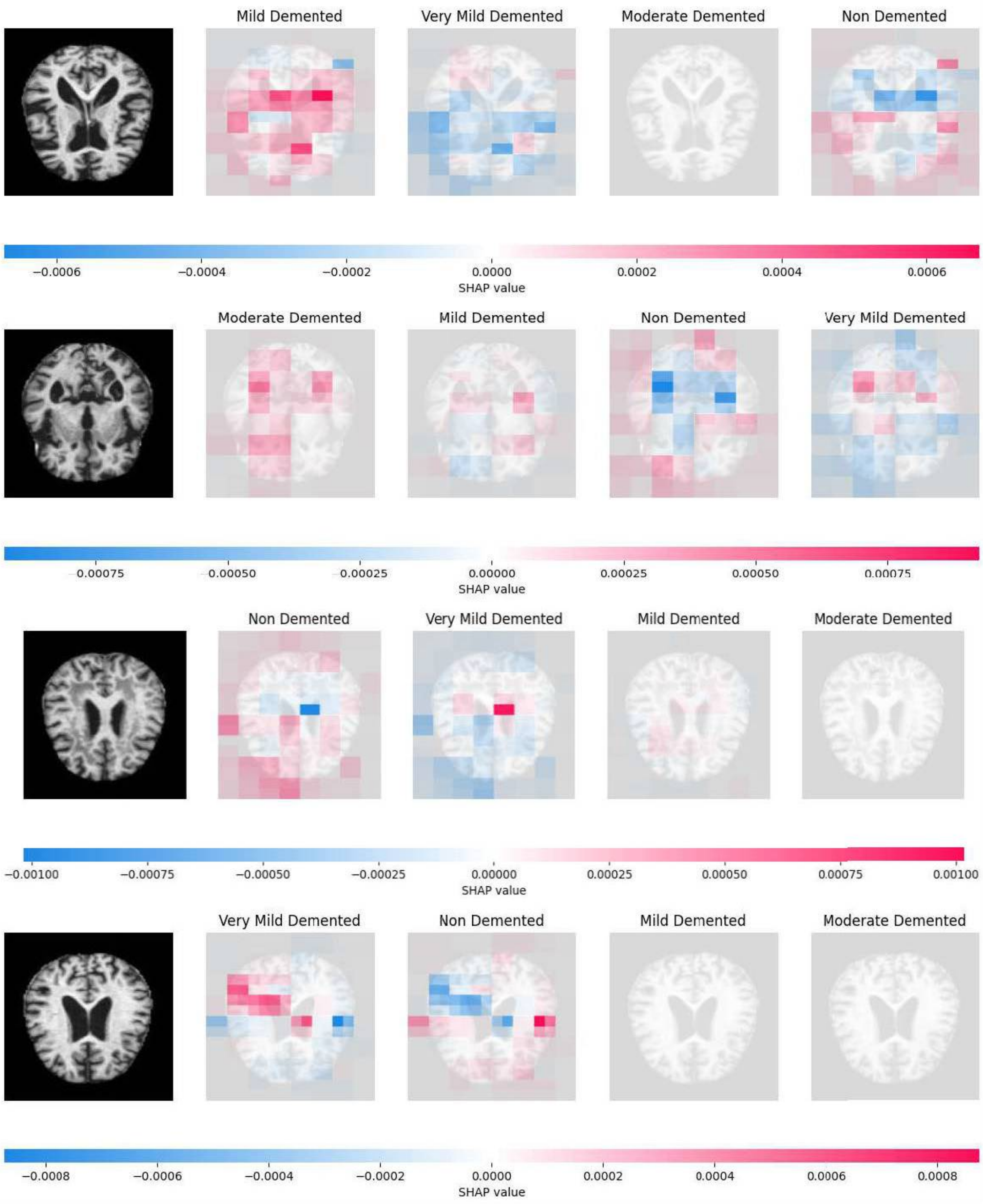


Attention map for dataset 1

Model	Optimizers	Precision(%)	Recall(%)	F1-score(%)	Specificity(%)	Cohen's Kappa	Time(Seconds)	Test accuracy(%)
ResNet50	**AdamW**	71.42	67.34	63.39	95.51	41.38	470.34	67.34
MobileNetV2	**AdamW**	95.48	95.47	95.45	97.52	92.39	450	95.47
VGG19	**AdamW**	90.11	89.53	89.34	97.75	82.24	670.19	89.53
Xception	**AdamW**	84.76	84.06	84.15	86.86	73.76	504.31	84.06
InceptionV3	**AdamW**	88.17	87.97	87.98	89.90	79.88	520.51	87.97
DenseNet121	**AdamW**	95.35	95.31	95.30	97.77	92.16	555.32	95.31
VGG16	**AdamW**	91.17	90.78	90.82	90.27	84.57	550.12	90.78
Proposed Model	**AdamW**	**99.53**	**99.53**	**99.53**	**99.76**	**99.23**	**450.98**	**99.53**

Comparing results of pre-trained CNN models with the proposed ViT-GRU on dataset 1

04



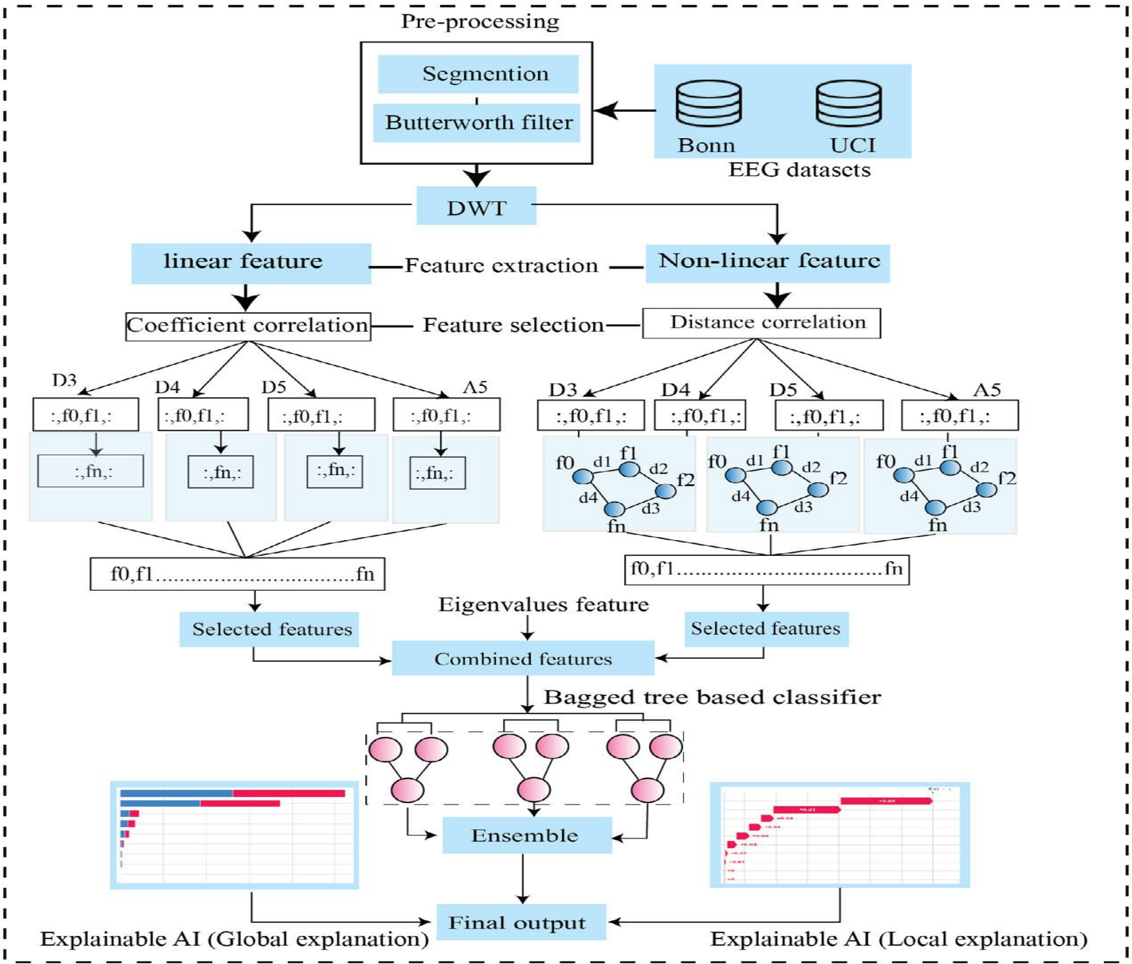
SHAP for dataset 1.

04

Study	Year	Model	Dataset	Image modalities	Number of class	Number of images	Performance	XAI
Al-Adhaileh et al. [58]	2022	AlexNet	AD dataset kaggle	MRI	—	—	Accuracy: 94.53% Precision:— Recall:— F1-score: 94.12%	Not used
Ullah et al. [59]	2022	CNN	Alzheimer MRI preprocessed dataset kaggle	MRI	4	6,400 samples increased to 12,800 using SMOTE	Accuracy: 99.38% Precision: 99% Recall: 99% F1-score : 99%	Not used
Biswas et al. [60]	2022	CNN	AD dataset kaggle	MRI	2	4800	Accuracy: 99.38% Precision: 99.70% Recall: 95% F1-score: 99.32%	Not used
Proposed study	**2023**	**ViT-GRU**	**Alzheimer MRI preprocessed dataset kaggle**	**MRI**	**4**	**6400**	**Accuracy: 99.53% Precision: 99.53% Recall: 99.53% F1-score: 99.53%**	**Used**
				MRI	**2**	**6400**	**Accuracy: 99.69% Precision: 99.69% Recall: 99.69% F1-score: 99.69%**	
			ADNI	**MRI**	**3**	**2970**	**Accuracy: 99.26% Precision: 99.27% Recall: 99.26% F1-score: 99.26%**	

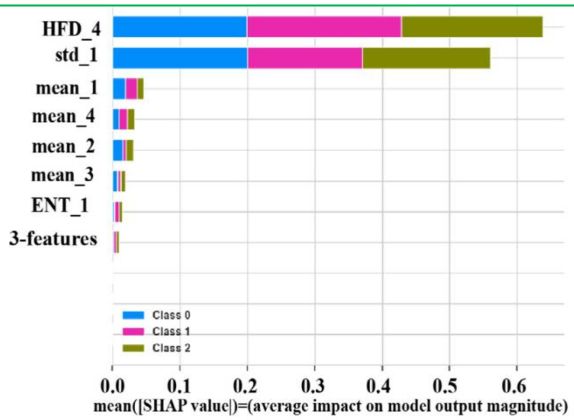
EEG Analysis

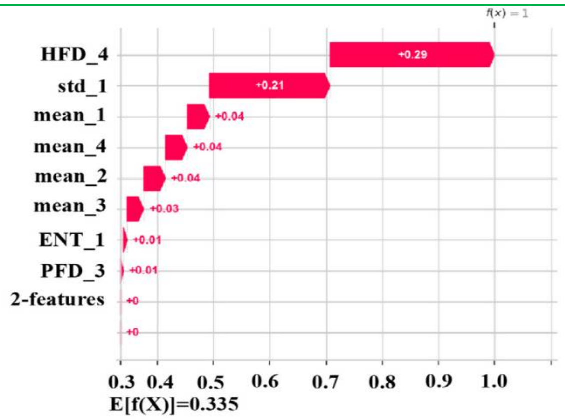
12



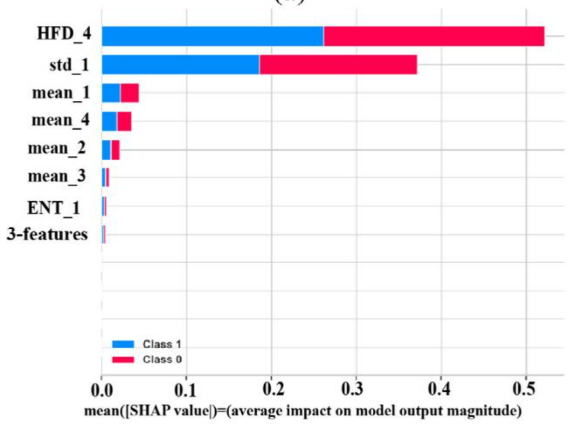
The block diagram of the proposed framework
for EEG epileptic seizure detection

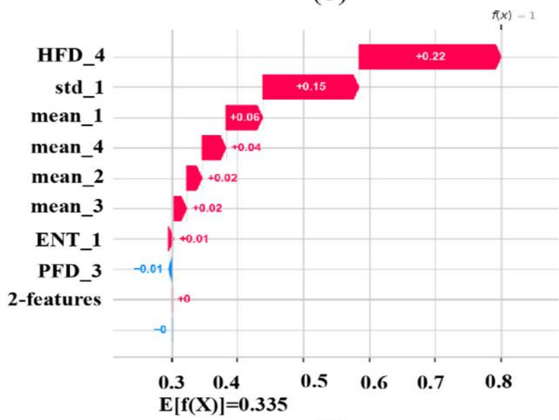
12





(a) (b)



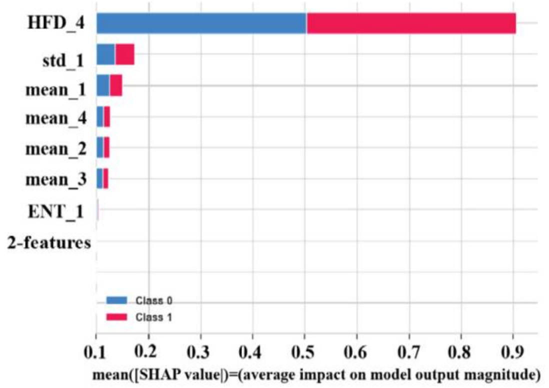


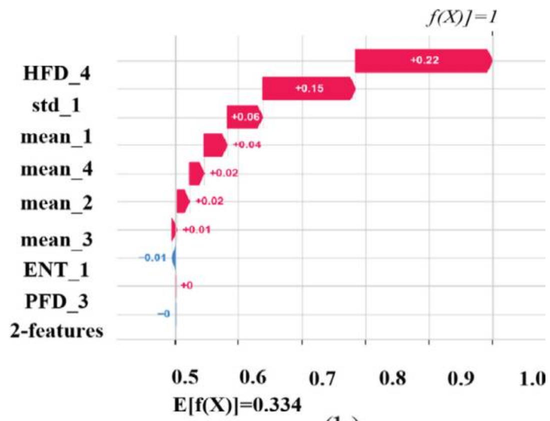
(c) (d)

Visual explanation for the propose model

- ✓ Based on different visualizations for binary and multi-class tasks using the Bonn EEG dataset. Subfigures (a) and (c) representsummary plots,
- ✓ While (b) and (d) shows waterfall plots

12



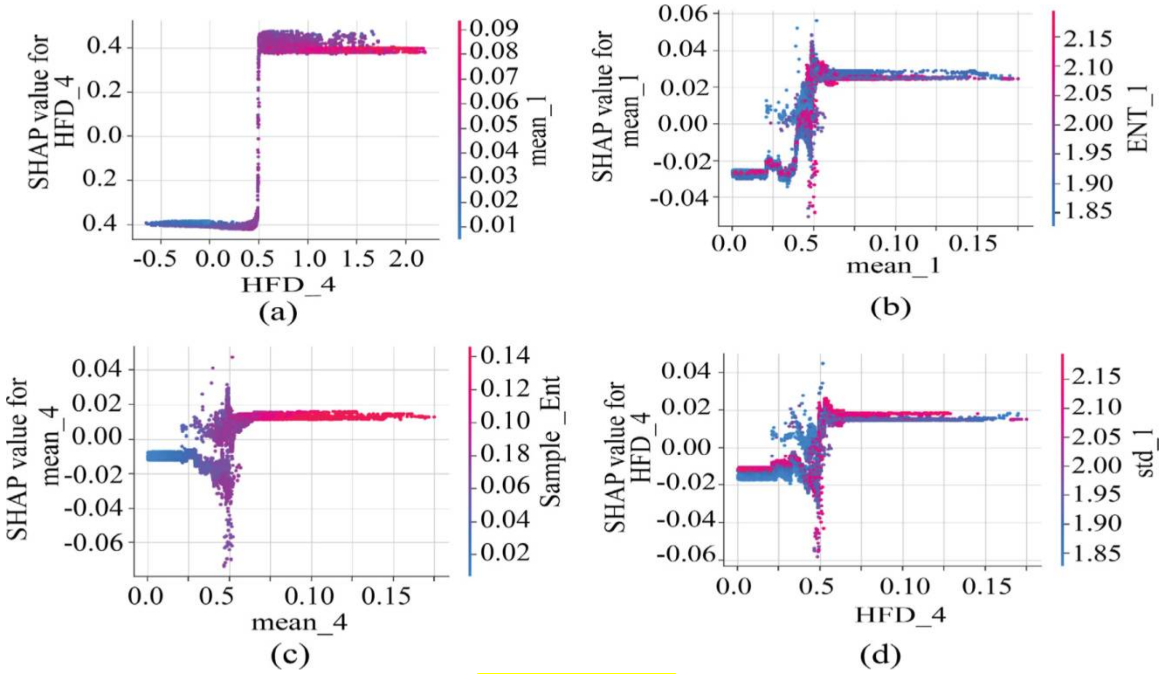


(a) (b)

The explainability of the proposed model

- ✓ Based on different visualizations for the binary class using the UCI- EEG dataset.
- ✓ Subfigure (a) represents the summary plot,
- ✓ While (b) depicts the waterfall plot

12

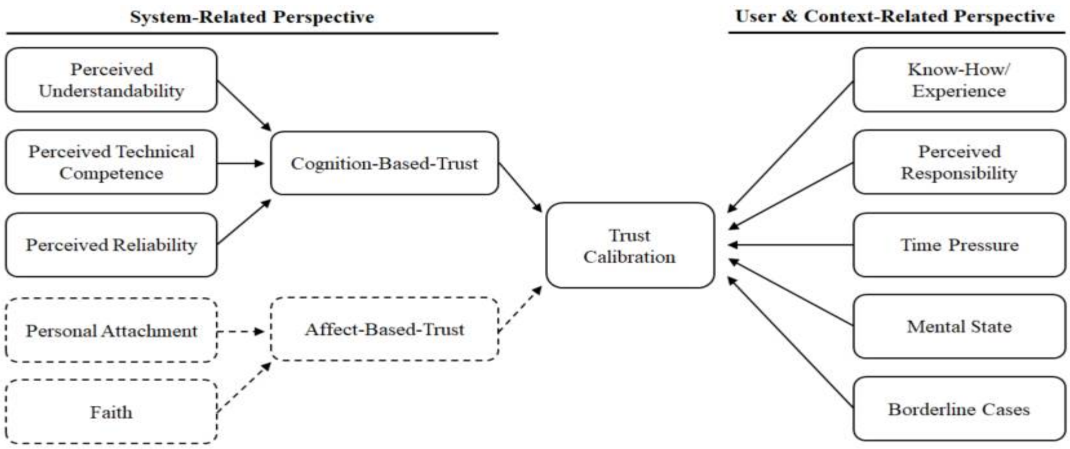


Dependency plot,

- ✓ Each subplot (a,b,c,d) represents a different features impact on the model output, with the color intensity indicating the features SHAP value
- ✓ Using different experiment type of the bonn and UCI EEG dataset

Health care

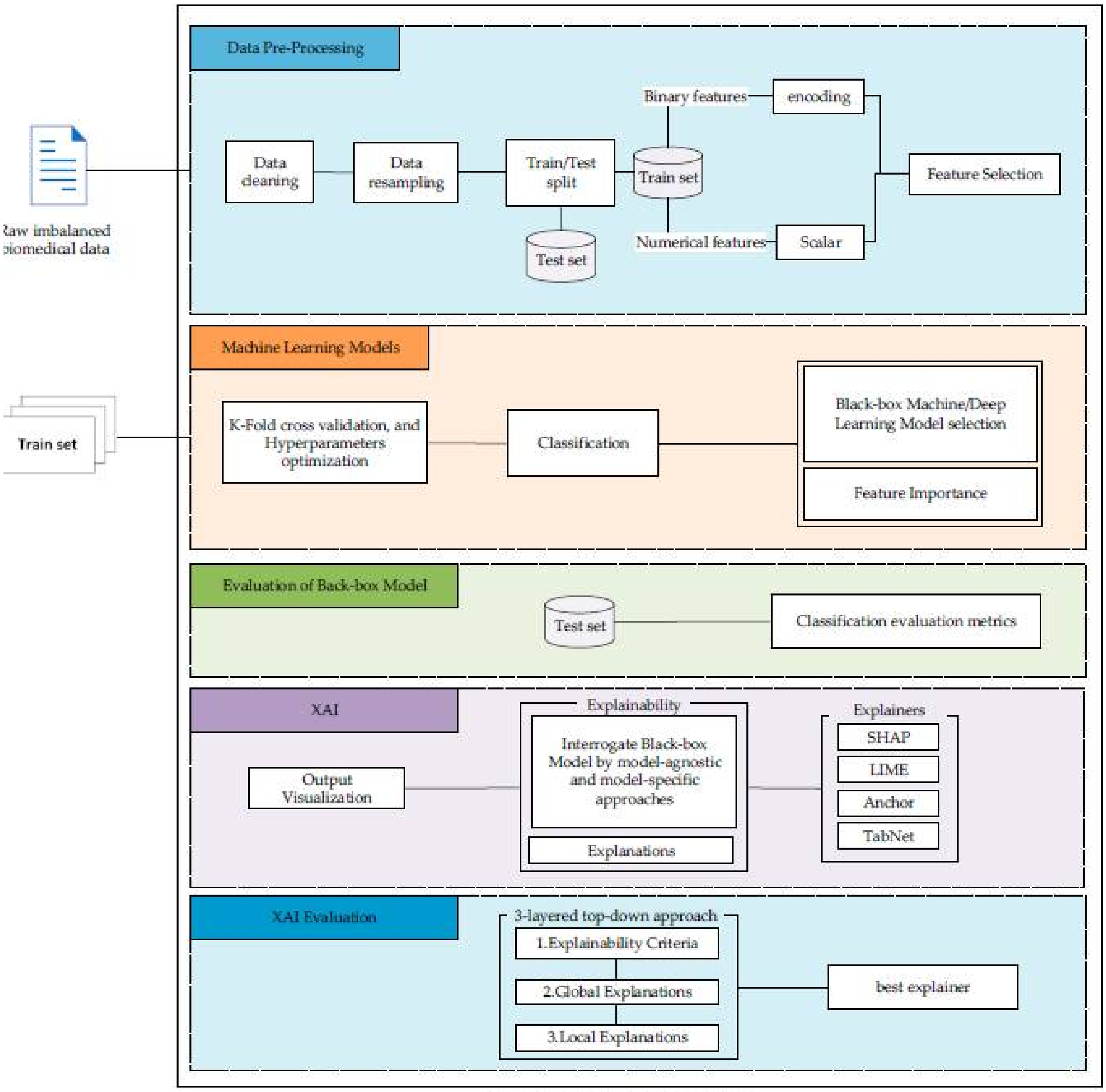
14



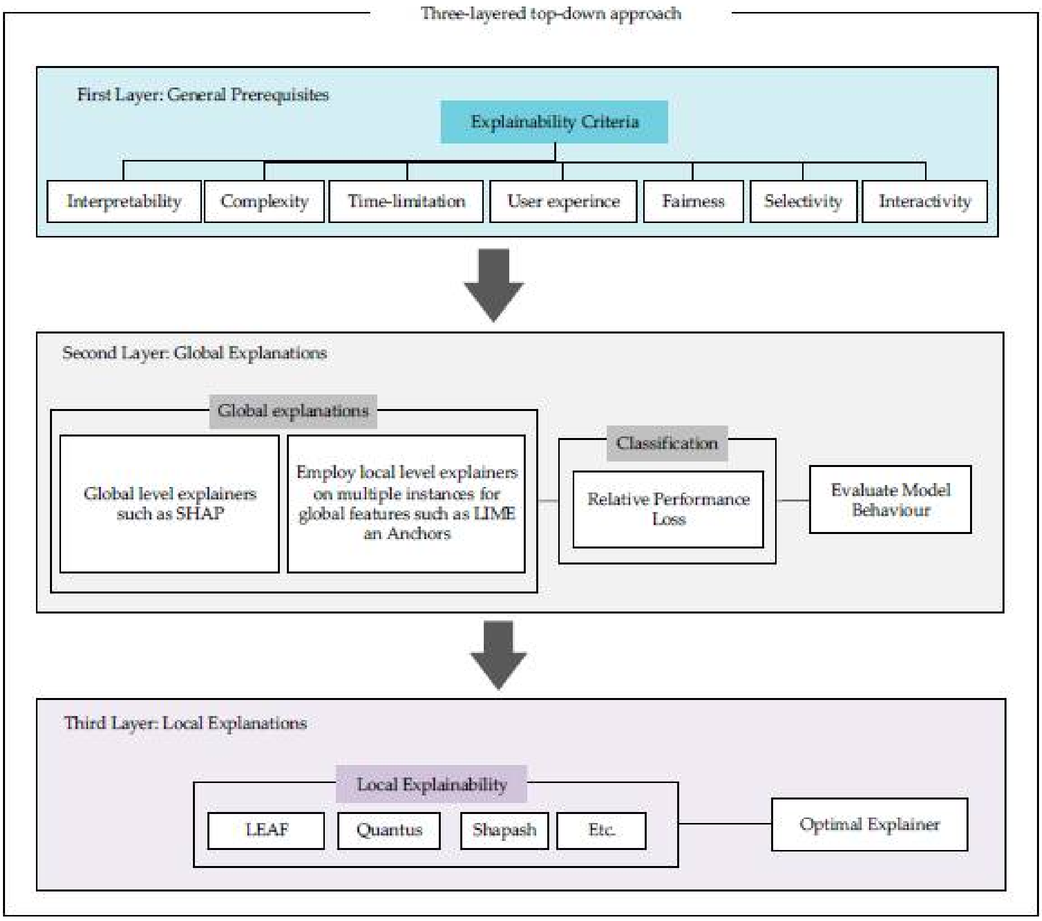
Trust calibration in relation to (X)AI in healthcare (based on HCTM,).

Heart (Cardiac) diseases

19



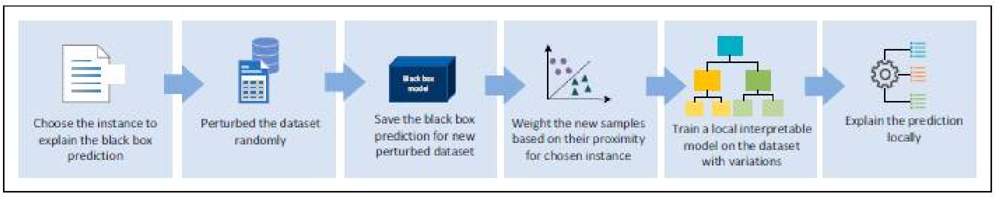
(a) – Methodological XAI Evaluation Process



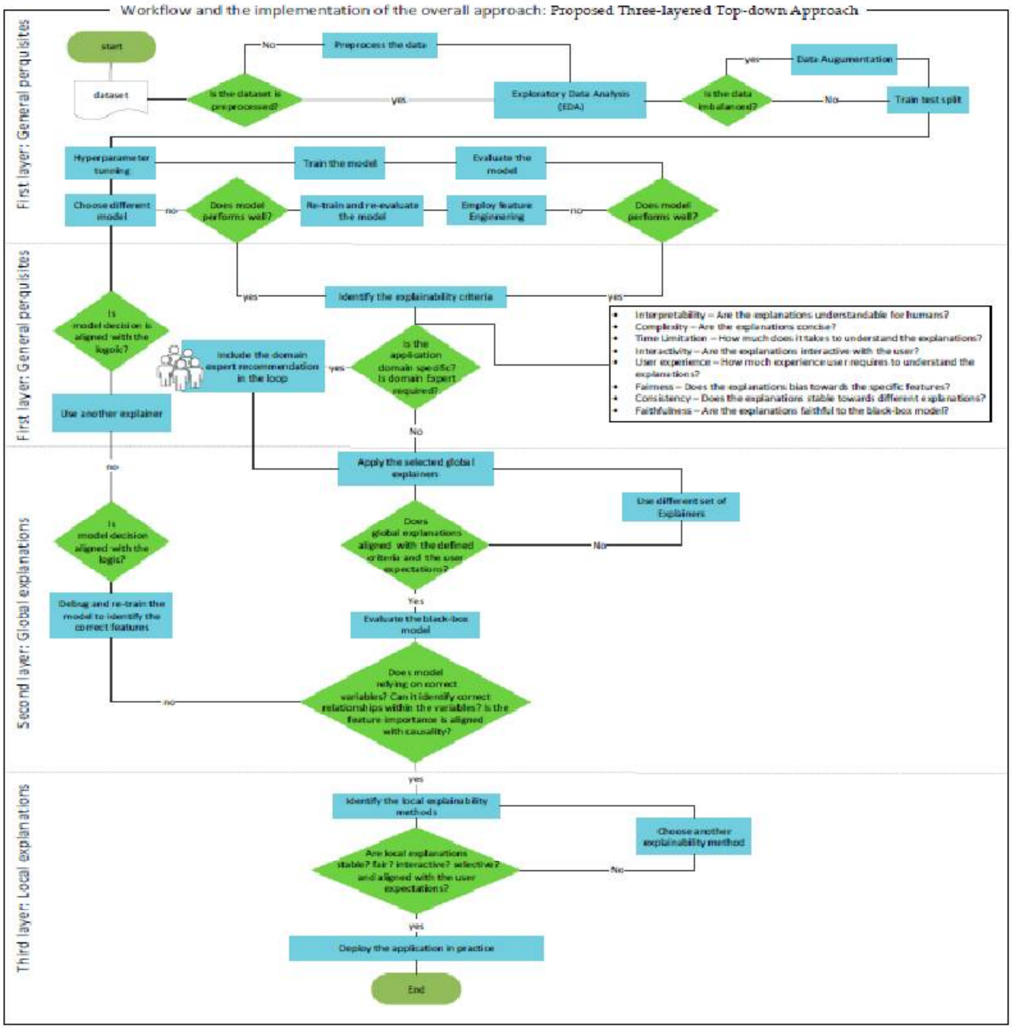
(b) – Proposed Three-layered Top-down Approach

Proposed Methodology

19



LIME Algorithm



Three-layer approach

19

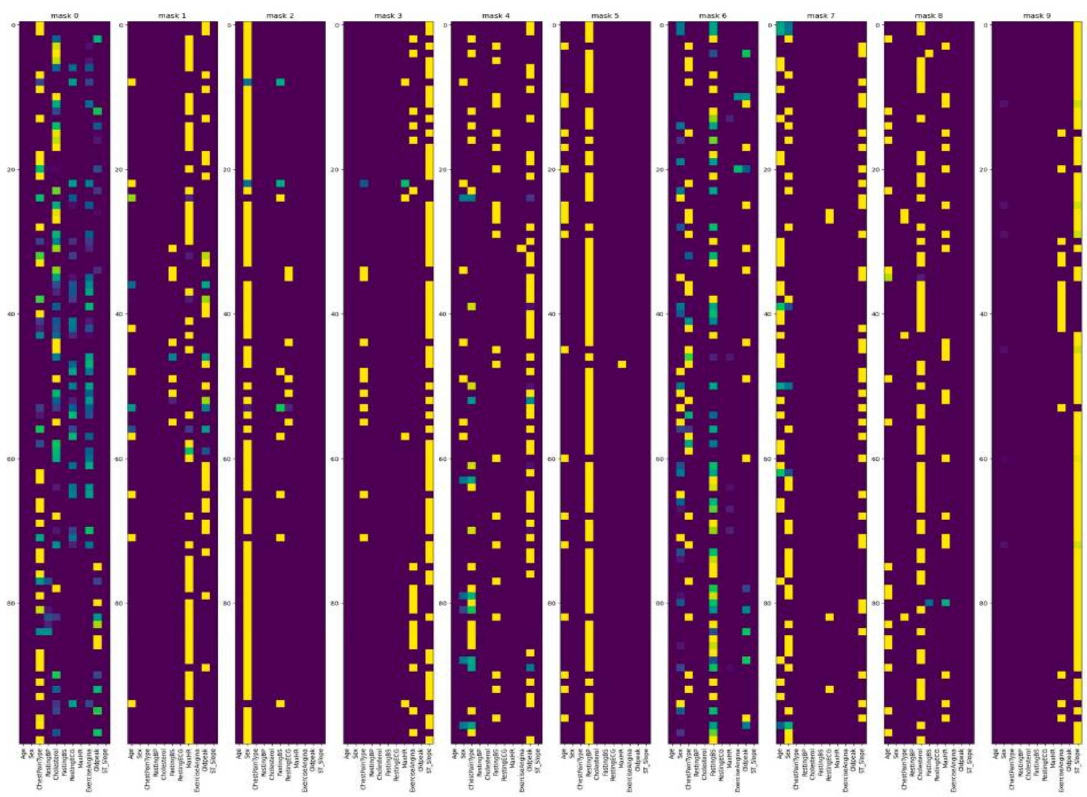
Feature Name	Description	Data Type
Age	In years between 28 and 77	Numerical
Sex	Gender coded as M for male and F for female	Categorical
Chest Pain Type	Type of the chest pain experienced by the patient during examination coded as TA: Typical Angina, ATA: Atypical Angina, NAP: Non-Anginal Pain, ASY: Asymptomatic	Categorical
RestingBP	Resting blood pressure in millimeters of mercury (mmHG) between 0 to 200	Numerical
Cholesterol	Serum cholesterol level of the patient in milligrams per deciliter (mg/dl) between 0 to 603	Numerical
FastingBS	Fasting blood sugar level coded as 1: if FastingBS > 120 mg/dL, and 0: otherwise	Categorical
RestingECG	Resting electrocardiogram results, coded as, Normal, ST: having ST-T waves abnormality, and LVH: showing probable or definite left ventricular hypertrophy	Categorical
MaxHR	Maximum heart rate achieved during exercise between 60 to 202	Numerical
Exercise Angina	Experienced angina during exercise which coded as Y: Yes, and N: No.	Categorical
Oldpeak	ST depression between −2.6 to 6.2	Numerical
ST_Slope	Slope of the peak exercise ST segment coded as, Up, Flat, and Down	Categorical
HeartDisease	Class label coded as 1 for heart disease and 0 for healthy.	Categorical

Heart Disease Dataset information

19

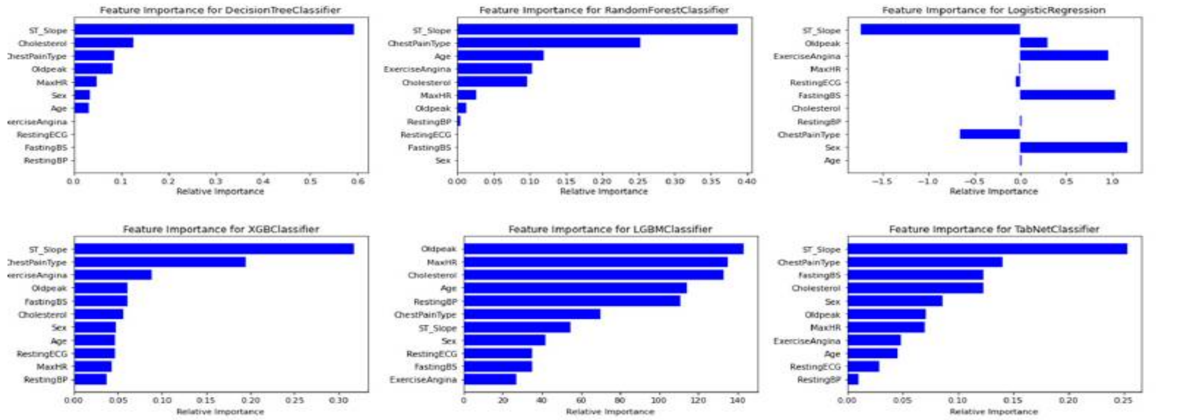
Heatmap for Heart Disease Classification

19



Tabnet Masks for Heart Disease Classification.

19



Model Intrinsic Feature Importance for Heart Disease Classification.

19

LIME Global Feature Importance for Heart Disease Classification

	LIME
DT	ST_Slope, Sex, ChestPainType
RF	ST_Slope, ExerciseAngina, Age
LR	ST_Slope, Sex, ExerciseAngina
XGboost	Sex, Cholesterol, ST_Slope
LightGBM	Sex, Cholesterol, ChestPainType
TabNet	Sex, FastingBS, ChestPainType

19

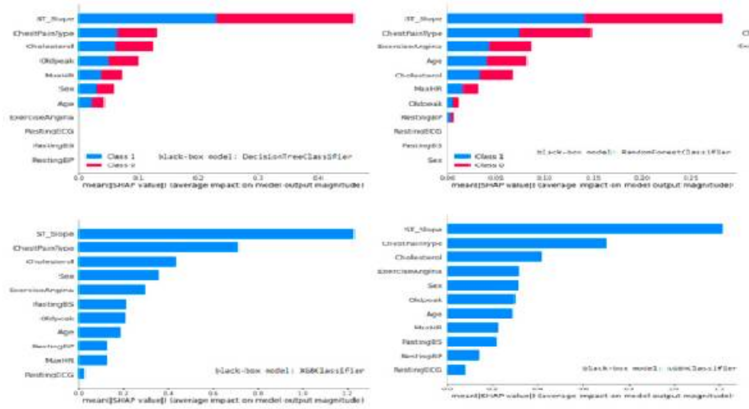


Figure 9. SHAP Global Feature Importance for Heart Disease Classification.

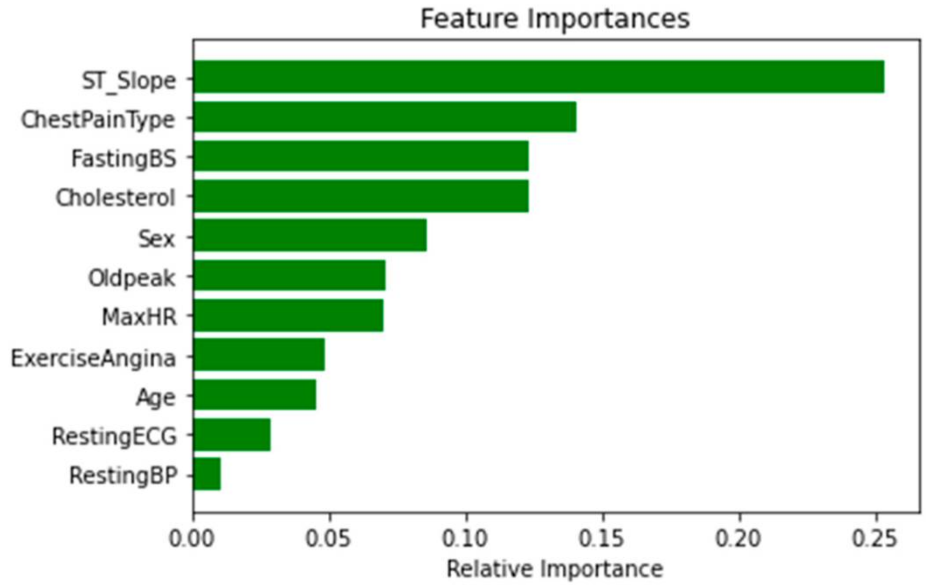
SHAP Global Feature Importance for Heart Disease Classification

19

Anchors Global Explanations for Heart Disease Classification

	Anchors
DT	ST_Slope, ChestPainType, Cholesterol
RF	ST_Slope, Age, ChestPainType
LR	ST_Slope, ExerciseAngina, ChestPainType
XGboost	ST_Slope, ChestPainType, Cholesterol
LightGBM	ST_Slope, ChestPainType, Oldpeak

19



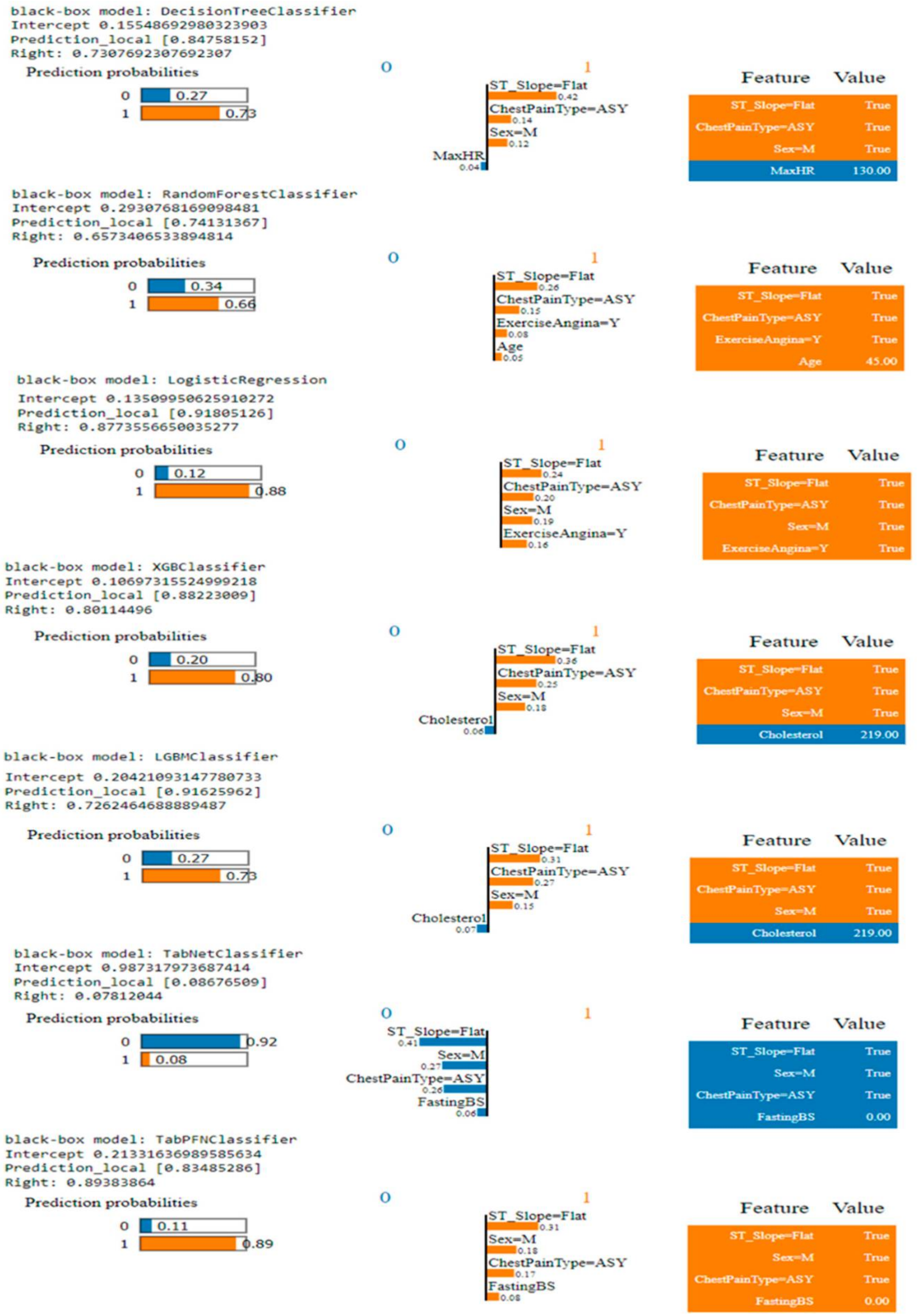
TabNet Global Feature Importance for Heart Disease Classification

19

Relative Performance Loss for Ensemble Trees
with SHAP, LIME, and Anchors
for Heart Disease Classification

	LIME	SHAP	Anchors
DT	10.63	25.53	25.53
RF	23.91	49.99	21.73
LR	43.90	46.34	46.34
XGBoot	60.60	81.81	81.81
LightGBM	27.77	38.88	58.33

19



LIME Local Explanations for Heart Disease Classification

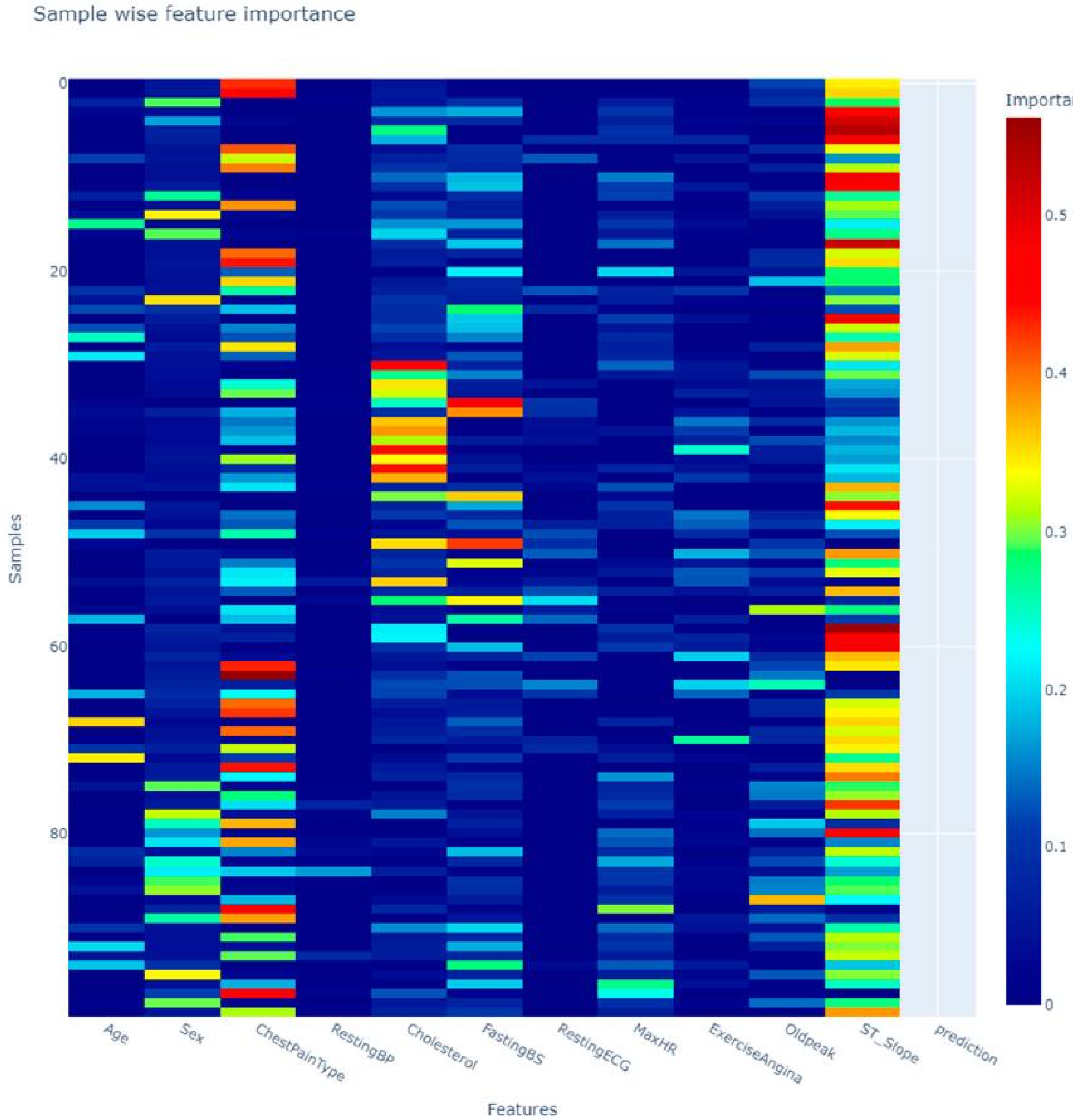
19

SHAP Local Explanations for Heart Disease Classification

19

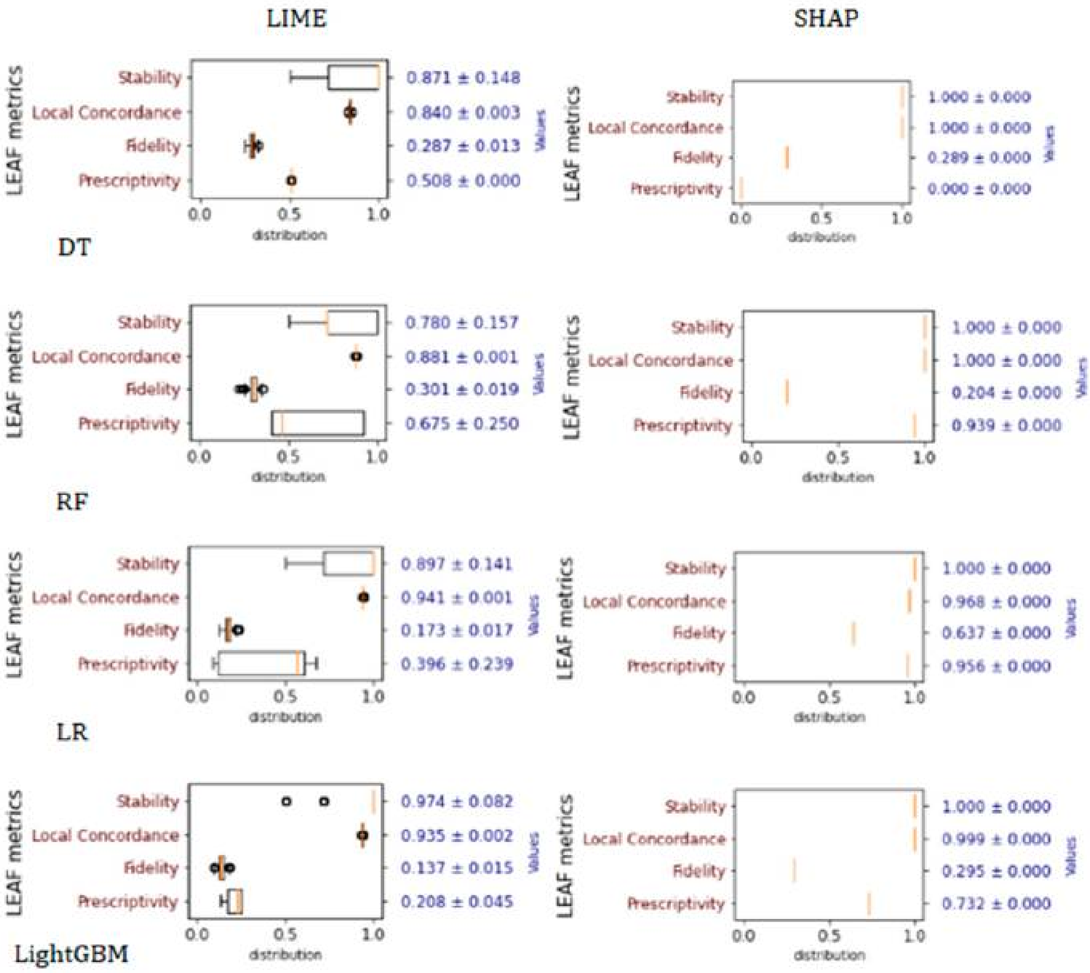
Anchors Local Explanations for Heart Disease Classification.

	Precision	Coverage	Anchor
DT	1.00	0.35	ST_Slope = Flat AND ChestPainType = ASY
RF	1.00	0.35	ST_Slope = Flat AND ChestPainType = ASY
LR	0.99	0.29	ST_Slope = Flat AND ExerciseAngina = Y
XGBoost	0.99	0.35	ST_Slope = Flat AND ChestPainType = ASY
LightGBM	0.97	0.35	ST_Slope = Flat AND ChestPainType = ASY
TabPFN	0.96	0.29	ST_Slope = Flat AND ExerciseAngina = Y



TabNet Local Explanations for Heart Disease Classification

19



LEAF evaluation framework for DT, RF, LR, and LightGBM for Heart Disease Classification

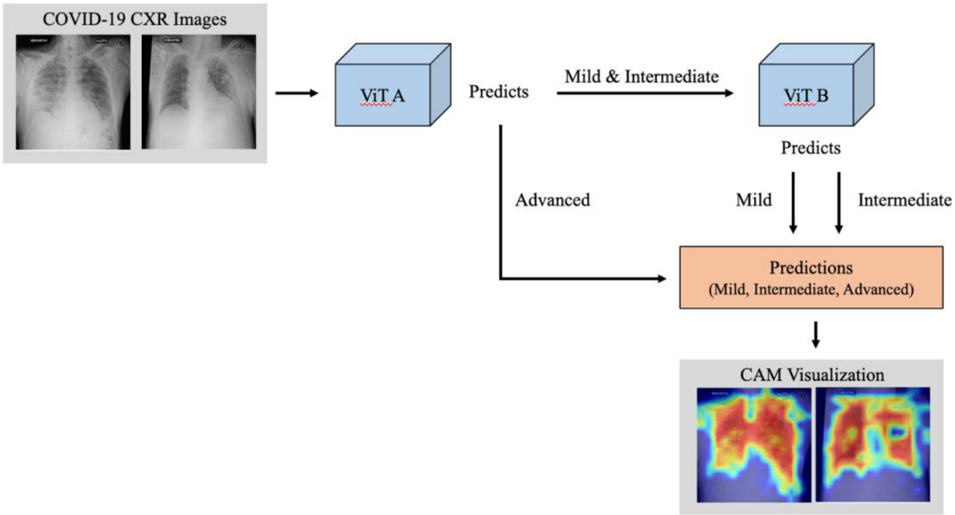
19

Classification report with 0 denoting a healthy person and 1 signifies diabetes

	Accuracy	Prec-0	Prec-1	Recall-0	Recall-1	F1-Score-0	F1-Score-1
DT	64.77	77	49	66	62	71	54
RF	72.15	83	58	73	70	78	63
LR	73.29	86	58	72	77	78	66
XGboost	73.86	81	61	79	63	80	62
LightGBM	70.45	79	59	76	60	77	58
TabNet	74.57	72	76	57	86	63	81
TabPFN	75.70	77	72	87	57	82	64

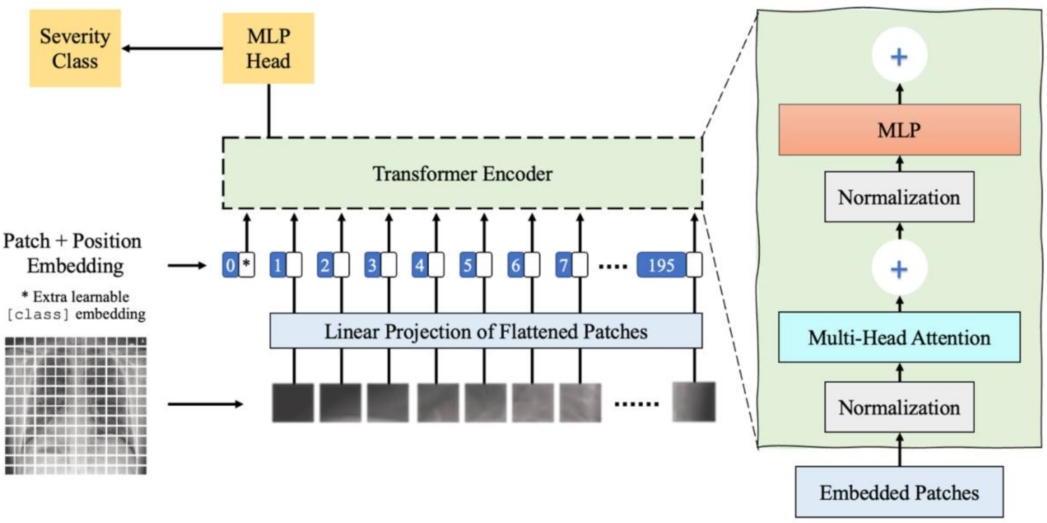
Pulmonary Disease

10



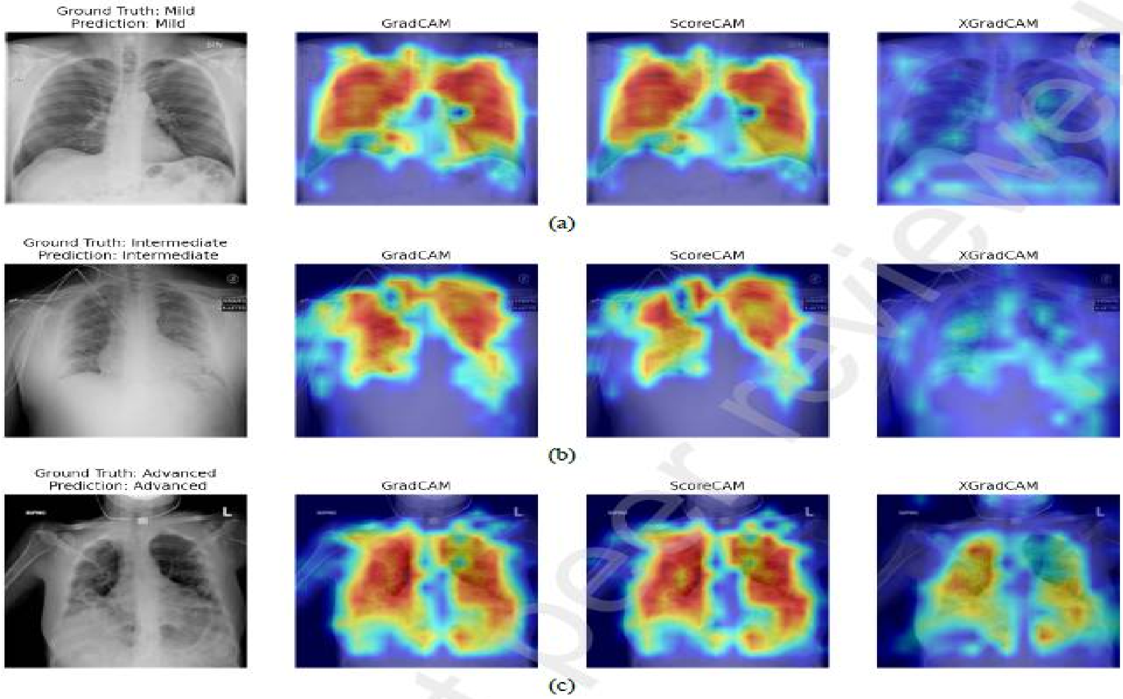
Proposed Architecture

10



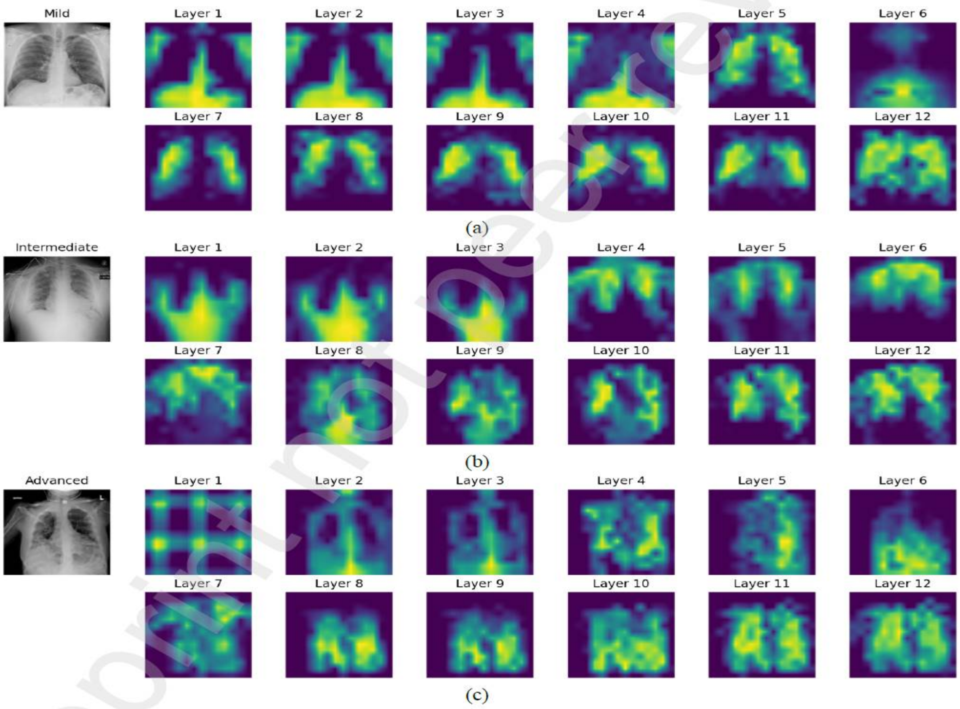
Overview of Vision Transformer architecture.

10



Visualizations obtained using
Grad-CAM, Score-CAM, and XGrad-CAM techniques for Mild, Intermediate, and Advanced cases

10



Grad-CAM visualizations on
Mild, Intermediate and Advanced cases across different layers (1 to 12) ofthe ViT

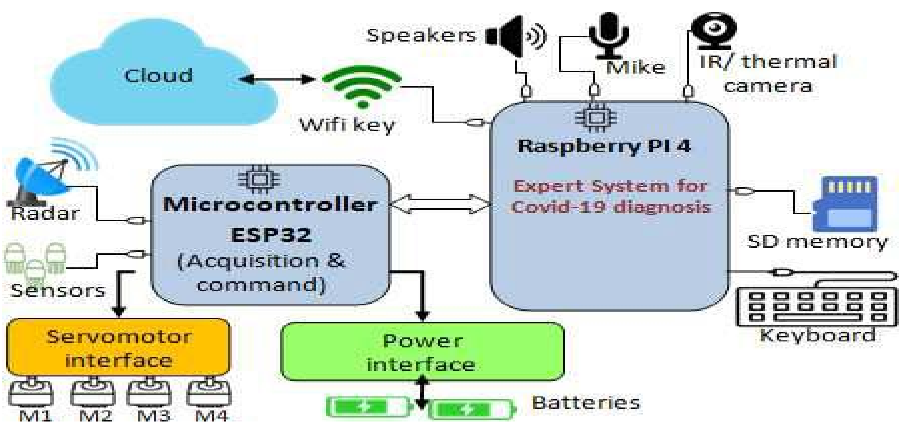
10

Models	Class Label	Precision	Recall	F1-Score	Overall ACC	I
		Raw				
ViT	Mild	0.8182	0.7759	0.7965	**0.7133**	
	Intermediate	0.6552	0.8143	0.7261		
	Advanced	0.6250	0.2273	0.3333		
	Macro Avg.	0.6995	0.6058	0.6186		
	Weighted Avg.	0.7138	0.7133	0.6957		
VGG16_bn	Mild	0.8936	0.7241	0.8000	0.6867	
	Intermediate	0.6538	0.7286	0.6892		
	Advanced	0.4000	0.4545	0.4255		
	Macro Avg.	0.6492	0.6358	0.6382		
	Weighted Avg.	0.7093	0.6867	0.6934		
ResNet50	Mild	0.8400	0.7241	0.7778	0.7067	
	Intermediate	0.6588	0.8000	0.7226		
	Advanced	0.5333	0.3636	0.4324		
	Macro Avg.	0.6774	0.6293	0.6443		
	Weighted Avg.	0.7105	0.7067	0.7014		

Comparison of performance between
ViT, VGG16_bn, and ResNet50 on the raw dataset

Covid-19 Disease

44



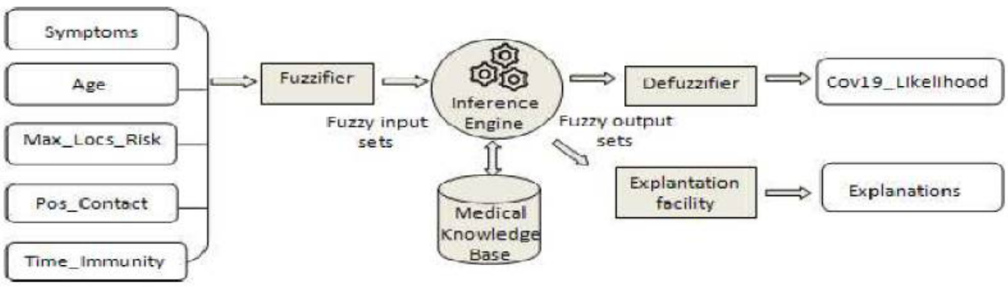
The hardware architecture for the diagnostic robot of COVID-19

44

Comparison between the studied expert systems for COVID-19 diagnosis

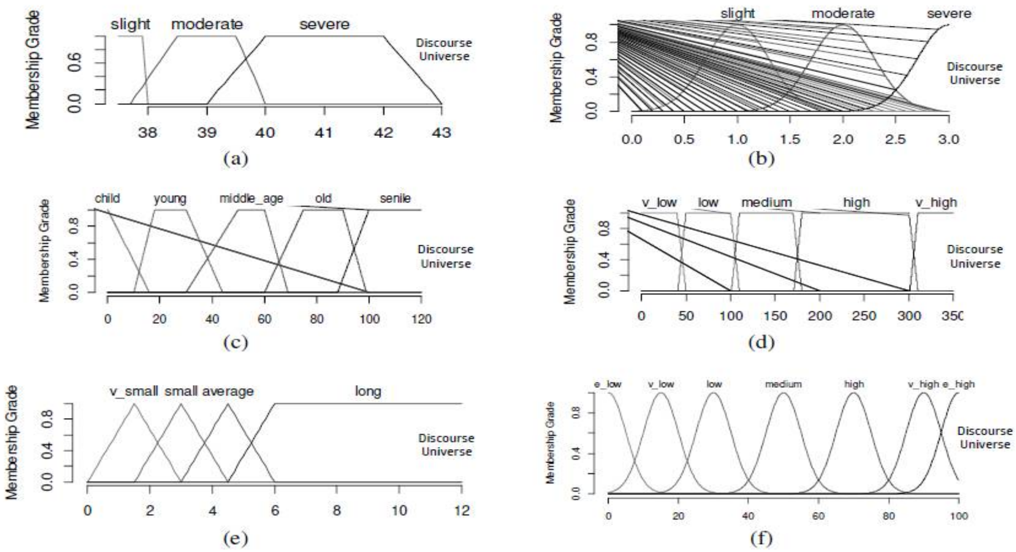
Diagnosis factors	Uncertainty support	Criteria Epidemic data update	Dynamic extensibility	Decision explicability	Tool
Symptoms contact history location history	No	No	No	No	python +CLIPS
Symptoms	Fuzzy logic	No	No	No	MATLAB toolbox
Symptoms, contact history, location history age	Certainty factors	No	No	No	Mobile App
Symptoms, measures	Triangular fuzzy numbers	No	No	No	Mobile and web apps
Symptoms hospitalization istory, epidemiological info, contact exposure	Fuzzy logic	No	Yes	Clinical rules + fuzzy sets plots	Web App
Symptoms, contact history, location history, age, immunity period (vaccine/infection)	Fuzzy logic	Yes	Yes (new rules, variables)	Yes (hybrid XAI)	Morfees-C19

44



Software architecture of MORFEES-C19 and its different components

44



Membership functions of the fuzzy variables:

- ✓ (a) fever, (b) headache,
- ✓ (c) age, (d) max risk, (e) time immunity
- ✓ (f) likelihood of contracting COVID-19

44

Table 2. Fuzzy variables of the proposal

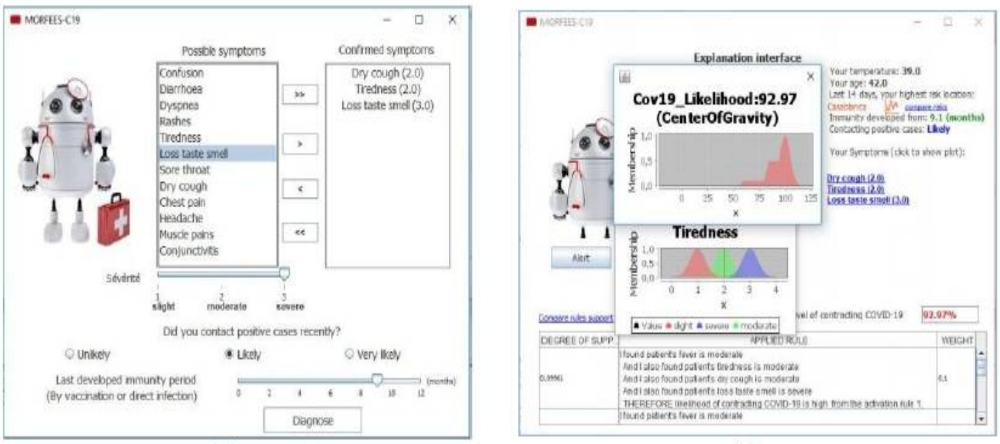
Linguistic variable	Linguistic terms	Universe of discourse
Fever	slight, moderate, and severe	From 36 to 43 ℃
The rest of symptoms	slight, moderate, and severe	from 0 to3
Age	child, young, middle-aged, old, and senile	greater than 0 years old
Max locs risk	very low, low, medium, high, and very high	greater than 1 case per 100K
Pos Contact	unlikely, likely, and very likely	from 0 to 3
Time Immunity	Very small, small, average, and long	from 0 to 12 months
Cov19_Likelihood	extremely low, very low, low, medium, high, very high, and extremely high	from 0% to 100%

44

RULE 1: IF Fever IS slight AND Tiredness IS slight AND Dry_cough IS slight AND Loss_taste_smell IS slight THEN Cov19_ Likelihood IS extremely low WITH 0.1;
RULE 2: IF Fever IS slight AND Tiredness IS slight AND Dry_cough IS slight AND Loss_taste_smell IS slight AND (Age IS old OR Age IS senile) THEN Cov19_Likelihood IS very low WITH 0.5;
RULE 3: IF Fever IS slight AND Tiredness IS slight AND Dry_cough IS slight AND Loss_taste_smell IS slight AND Time_ Immunity IS long THEN Cov19_ Likelihood IS very low WITH 0.5;
RULE 4: IF Fever IS slight AND Tiredness IS slight AND Dry_cough IS slight AND Loss_taste_smell IS slight AND (Pos_Contact IS likely OR Pos_ Contact IS very likely) THEN Cov19_ Likelihood IS low;
RULE 5: IF Fever IS slight AND Tiredness IS slight AND Dry_cough IS slight AND Loss_taste_smell IS slight AND (Max_Locs_ Risk IS high OR Max_Locs_ Risk IS very high) THEN Cov19_Likelihood IS low;
RULE 6: IF Fever IS slight AND Tiredness IS slight AND Dry_cough IS slight AND Loss_taste_smell IS slight AND (Diarrhoea IS NOT slight OR Conjunctivitis IS NOT slight OR Headache IS NOT slight OR Muscle pains IS NOT slight OR Sore throat IS NOT slight OR Rashes IS NOT slight) THEN Cov19 Likelihood IS very low WITH 0.5;
RULE 7: IF Fever IS slight AND Tiredness IS slight AND Dry_cough IS slight AND Loss_taste_smell IS slight AND (Confusion IS NOT slight OR Chest_ pain IS NOT slight OR Dyspnea IS NOT slight) THEN Cov19_Likelihood IS low;
RULE 8: IF Fever IS NOT slight AND Tiredness IS slight AND Dry_cough IS slight AND Loss_taste_smell IS slight THEN Cov19_ Likelihood IS very low WITH 0.1;
RULE 9: IF Fever IS NOT slight AND Tiredness IS NOT slight AND Dry_cough IS slight AND Loss_taste_smell IS slight THEN Cov19_ Likelihood IS low WITH 0.1;
RULE 10: IF Fever IS NOT slight AND Tiredness IS NOT slight AND Dry cough IS NOT slight AND Loss taste smell IS slight THEN Cov19_ Likelihood IS medium WITH 0.1;

The sample defined fuzzy rules

44

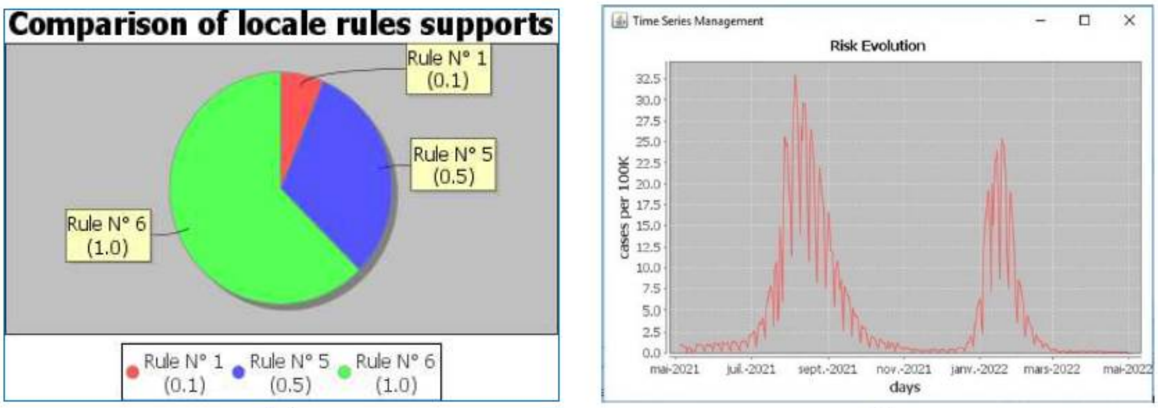


(a) (b)

Input and output frames of MORFEES-C19

- ✓ User input frame and (b) output frame with explanations

44



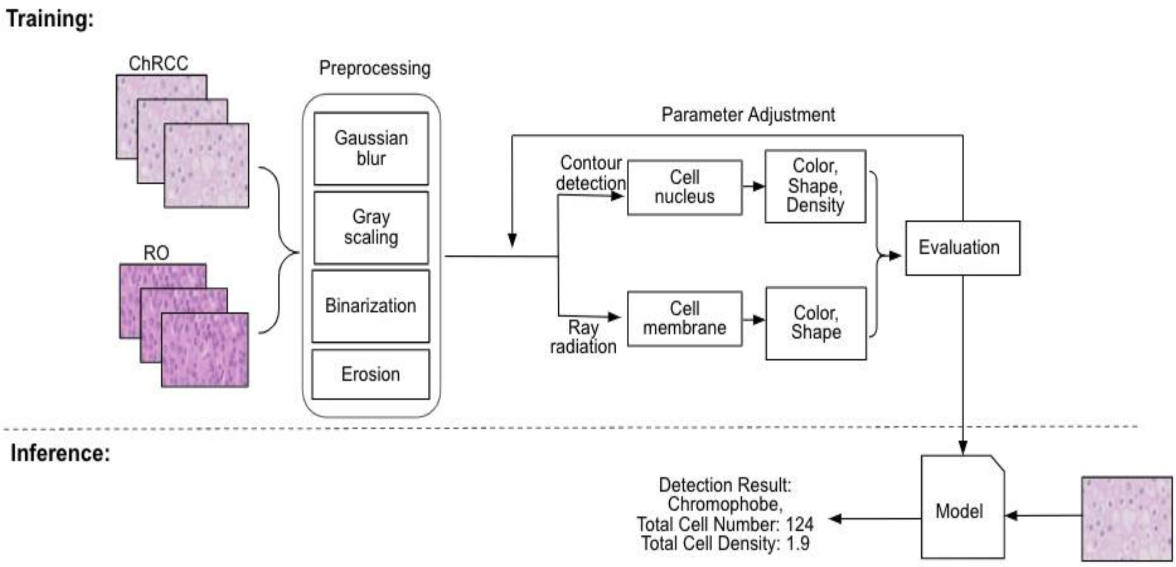
(a) (b)

Visual explanations

- ✓ (a) Pie chart of rules supports and
- ✓ (b) Time series of risk evolution

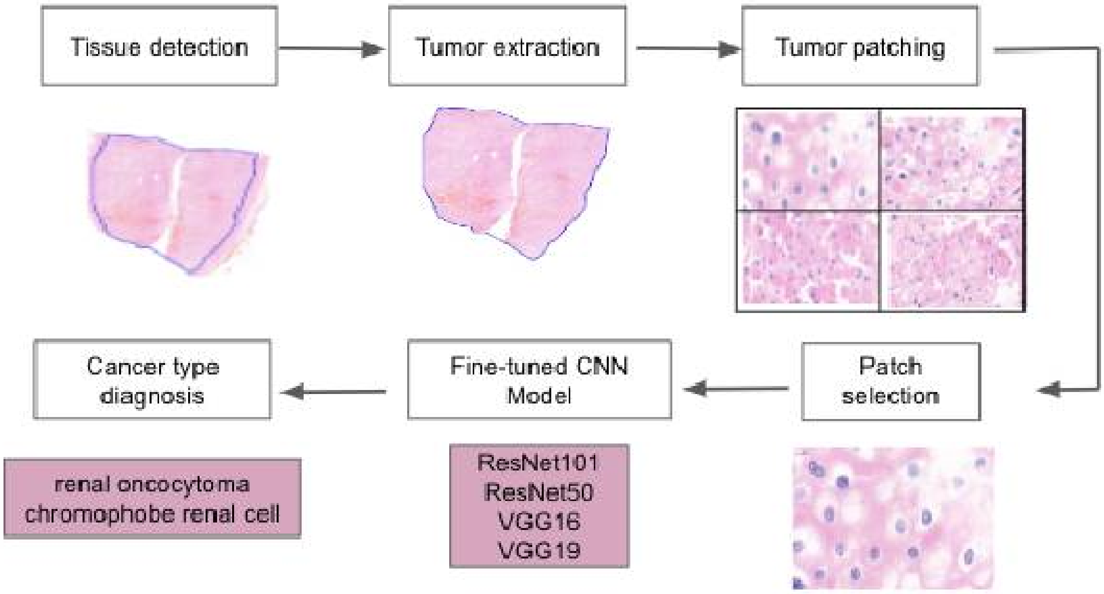
Diabetes Disease

25



Process flow of proposed framework
for RO and ChRCC image classification.

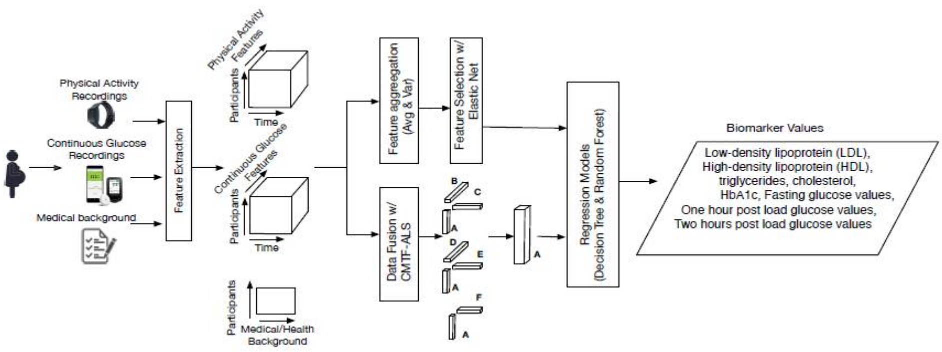
25



Workflow of analysis with neural network.

- ✓ This figure presents a schematic representation of the patching and
- ✓ Classification process used
- ✓ To differentiate between chromophobe and oncocytoma cells
- ✓ In kidney tumor samples

30



Framework for early prediction of biomarker values associated with the presence of GDM

Key components:

- ✓ Data collection as the initial step;
- ✓ Feature extraction,
- ✓ Capturing acute, cumulative, and magnitude changes;
- ✓ Feature aggregation and selection,
- ✓ Or data fusion with coupled matrix tensor factorisation-alternating least squares;
- ✓ And concluding with tree-based regression models
- ✓ For effective continuous biomarker value prediction

30

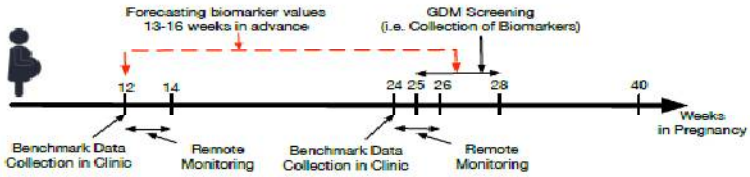
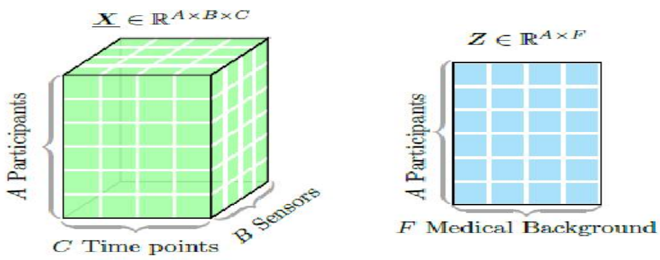


Illustration of benchmark data collection

- ✓ At 12 and 24 weeks,
- ✓ Remote monitoring in the first round between weeks 12-14,
- ✓ And in the second round between weeks 24-26, along with gdm screening
- ✓ Between week 25 and 28 during pregnancy at a south African Antenatal clinic.
- ✓ Dashed red line represents the forecasting period
- ✓ For the biomarker values collected between week 25 and 28, based on the benchmark data collected in week 12

30

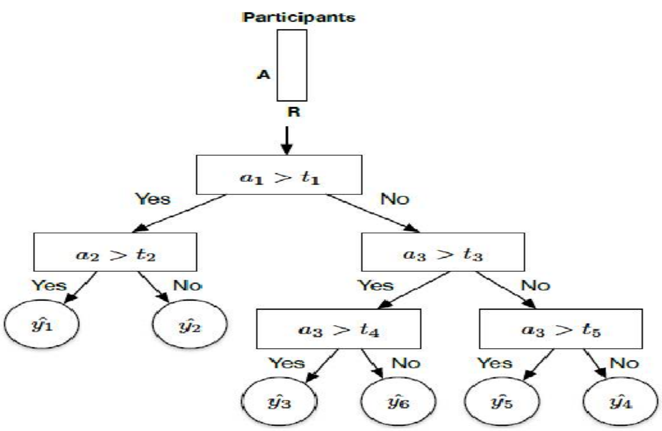


Coupled Matrix - Tensor example:

Tensors often share one or more modes

- ✓ i.e. A :participants
- ✓ X : Activity monitoring sensor recording tensor and
- ✓ Z : medical background matrix.
- ✓ As the vertical line indicates, these two datasets are coupled in the "participants" dimension

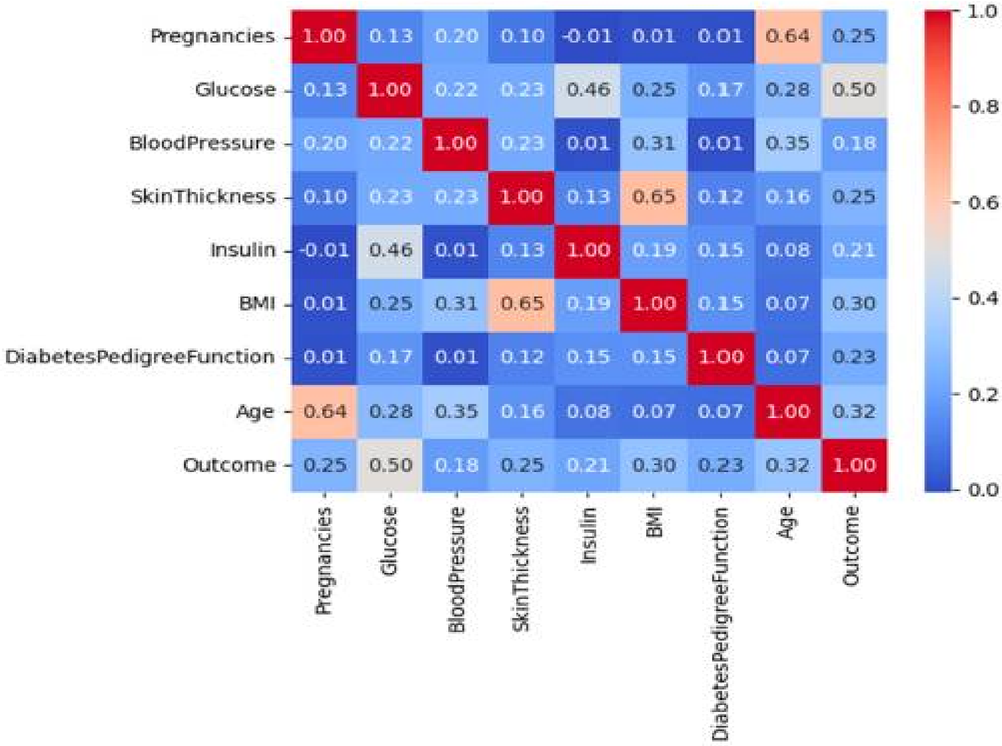
30



Decision tree for regression

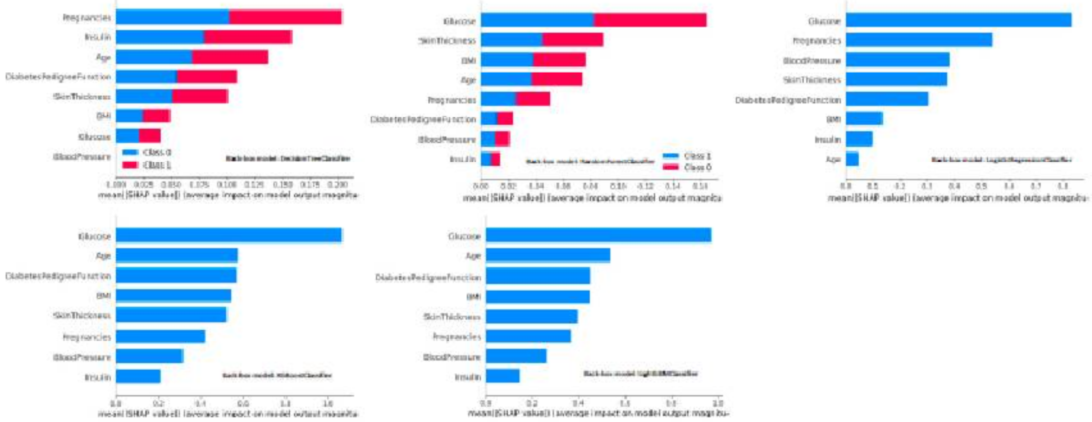
- ✓ Utilized to predict biomarker values in the experiments

19



Heatmap for Diabetes Disease Classification

19



SHAP Global Feature Importance for Diabetes Disease Classification.

19

Relative Performance Loss for Ensemble Trees
with SHAP, LIME, and Anchors
for Diabetes
Disease Classification.

	LIME	SHAP	Anchors
DT	6.89	−29.31	5.17
RF	−3.7	14.81	16.66
LR	9.8	9.80	1.96
XGBoot	46.51	32.55	46.51
LightGBM	−3.84	11.53	19.23

Relative Performance Loss for Deep Learning models for Diabetes Disease Classification

Deep Learning Model	Log Loss
TabNet	39.95
TabPFN	25.72

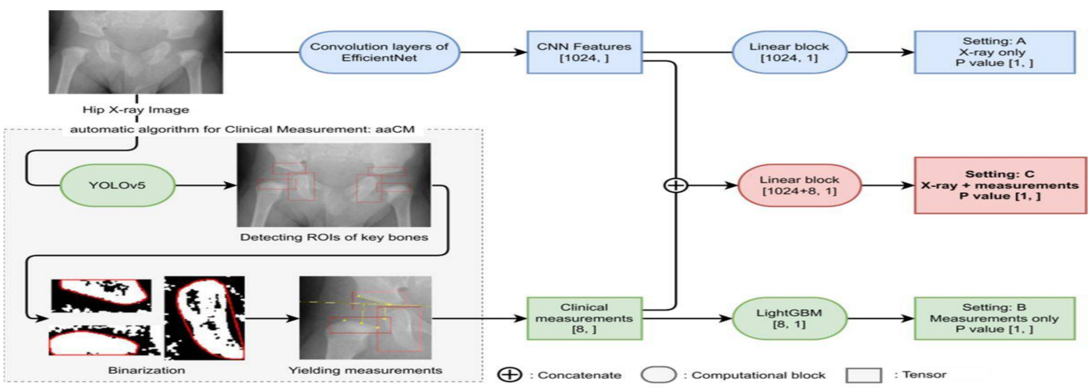
19

Anchors Local Explanations for Diabetes Disease Classification

	Precision	Coverage	Anchor
DT	097	0.09	Age > 28 AND Insulin > low 165.00 AND Glucose > 143.25
RF	0.91	0.02	Glucose > 143.25 AND Age > 28.00 AND Insulin > 165.00 AND BMI > 28.40 AND DiabetesPedigreeFunction > 0.41 AND Pregnancies > 2.00 AND BloodPressure ≤ 72.00
LR	0.98	0.01	Glucose > 143.25 AND Pregnancies > 2.00 AND DiabetesPedigreeFunction > 0.41 AND BloodPressure ≤ 64.00
XGBoost	0.95	0.02	Glucose > 143.25 AND Age > 28.00 AND BMI > 28.40 AND BloodPressure ≤ 64.00
LightGBM	0.99	0.01	Glucose > 143.25 AND Age > 28.00 AND DiabetesPedigreeFunction > 0.41 AND BMI > 28.40 AND BloodPressure ≤ 64.00
TabPFN	0.98	0.09	Glucose > 143.25 AND Age > 28.00 AND BMI > 28.40 AND DiabetesPedigreeFunction > 0.41

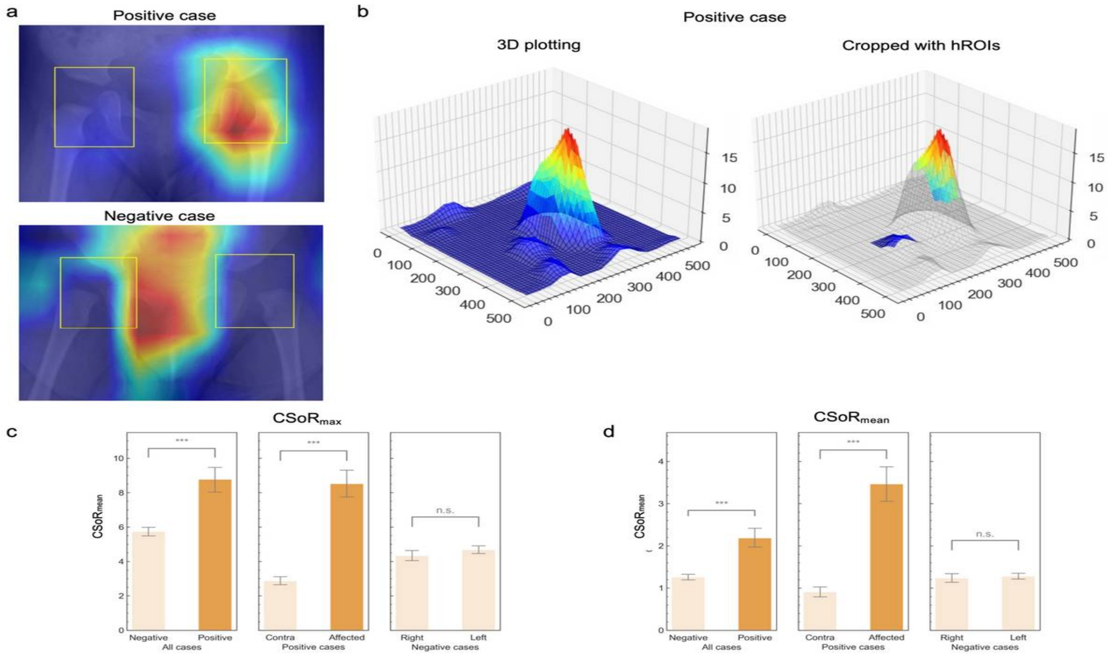
Orthopedic diseases

32



Overview of end-to-end models

32



Evaluation by GSoR

a) Representative images of positive and negative cases in which Grad-CAM heat-map was integratedwith hROIs.

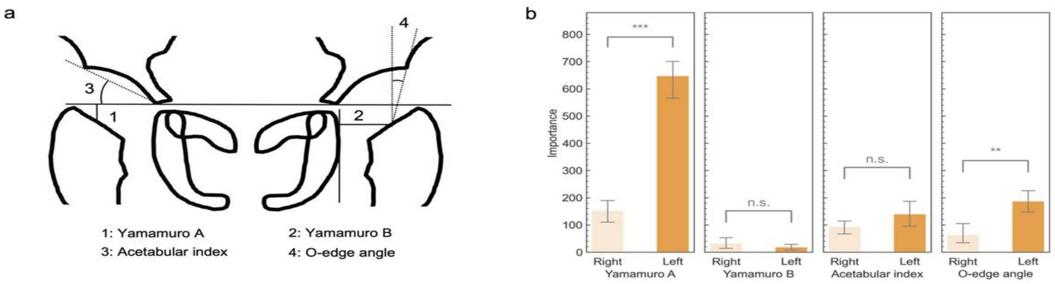
b) 3D plotting of relative CAM activities (left) and cropping by hROIs (right).

c) CSoRmax of positive and negative cases (left), the affected side and contralateral side in the positive cases(middle),and the right and left side in the negative case(right).

d) CSoRmean of the positive and negative cases(left), the affected side and contralateral side in the positive cases (middle), and the right and left side inthe negative case.

A non-paired student t-test was used. n.s. P>0.05, *P<0.05, **P<0.01, ***P<0.001.Contra.: Contralateral

32

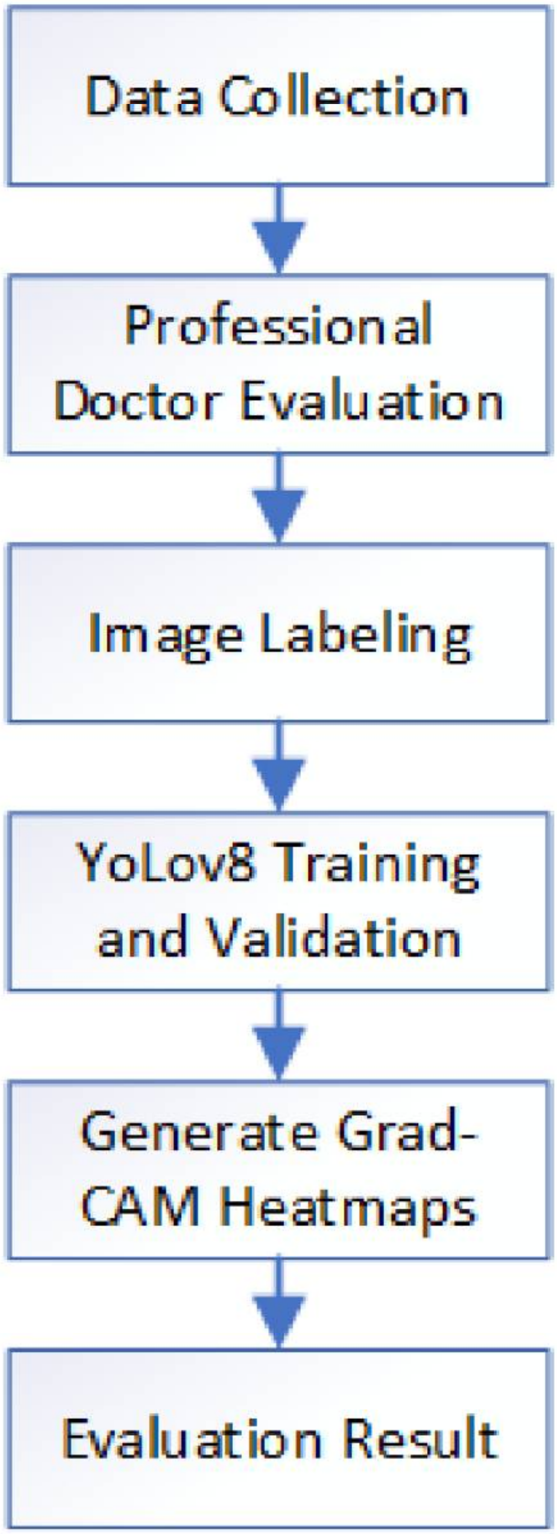


Feature Importance scores on the parameters

a) Scheme of radiographic parameters.

b) Feature importance scores of Yamamuro A, Yamamuro B,acetabular index, and O-edge angle on left and right sides.

- ✓ A non-paired student t-test was performed for comparison. n.s. P>0.05, *P<0.05, **P<0.01, ***P<0.001

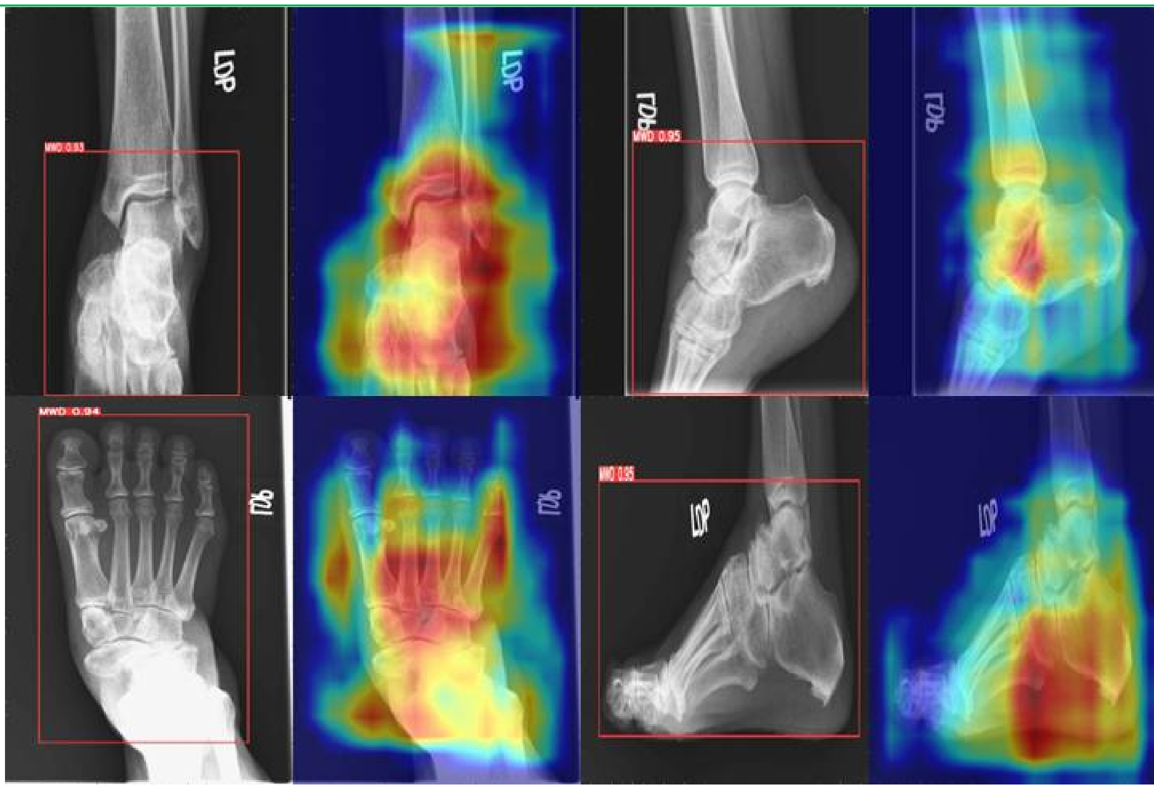


Research process.

20

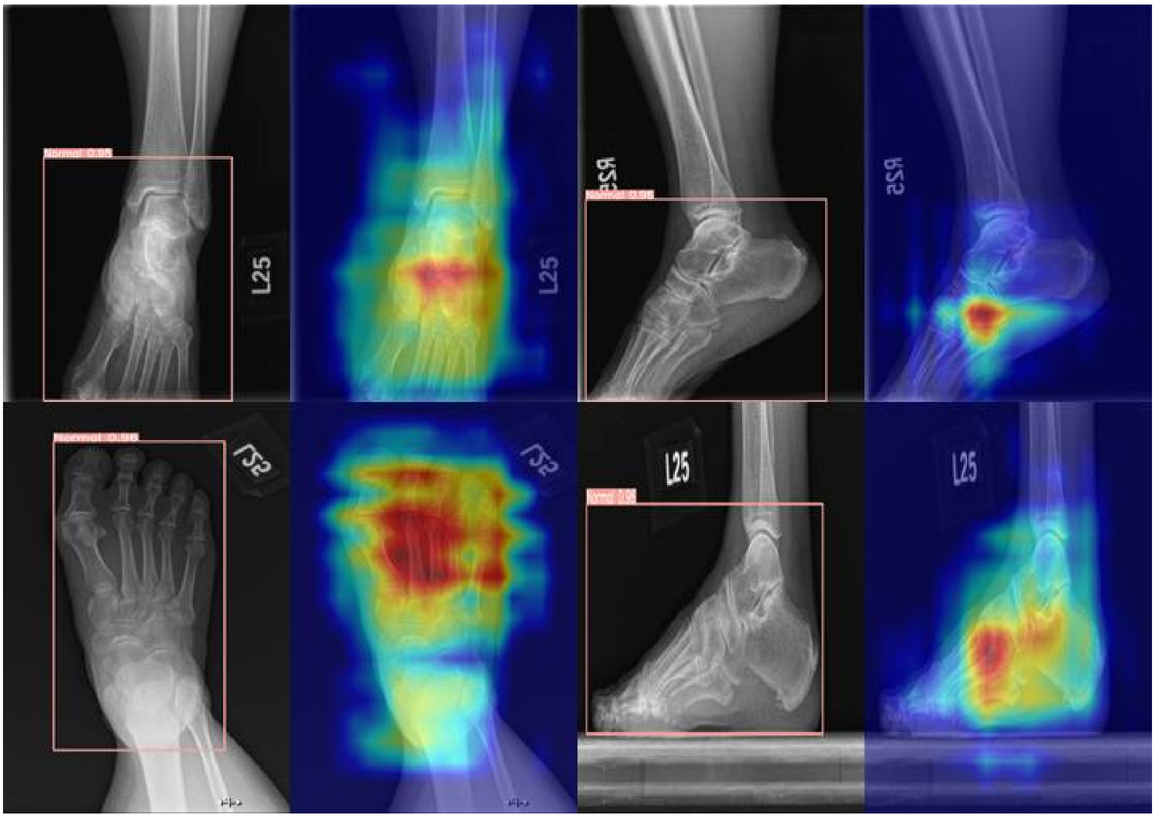
Orthopedic imaging perspective

20



Heatmaps of detection of MWD diseases

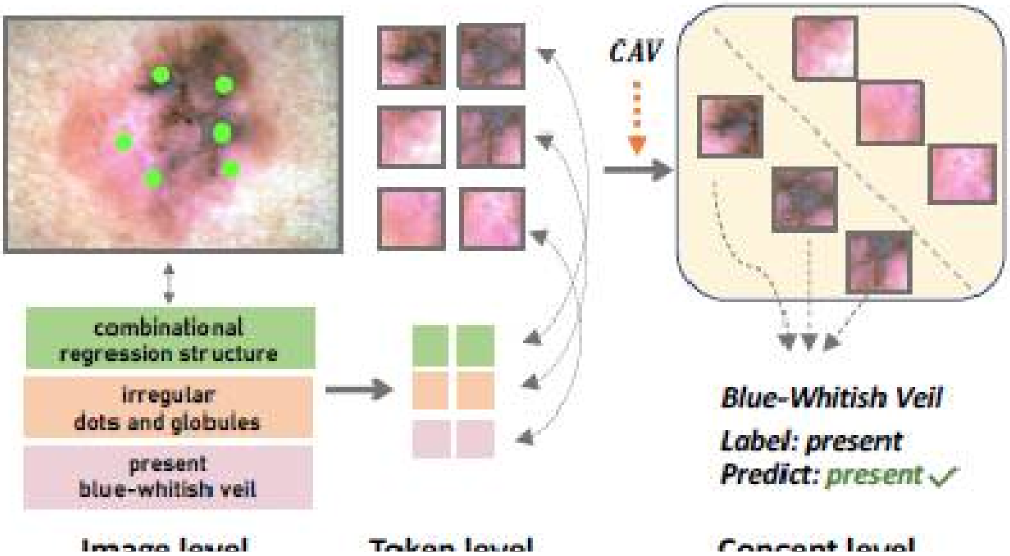
20



Heatmaps of detection of Normal diseases

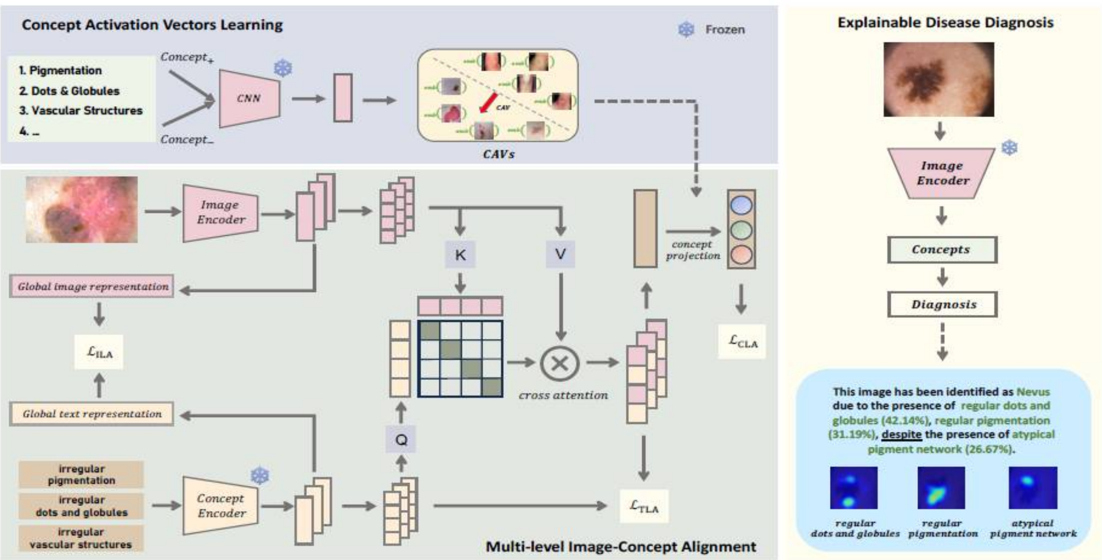
Dermatology diseases

21



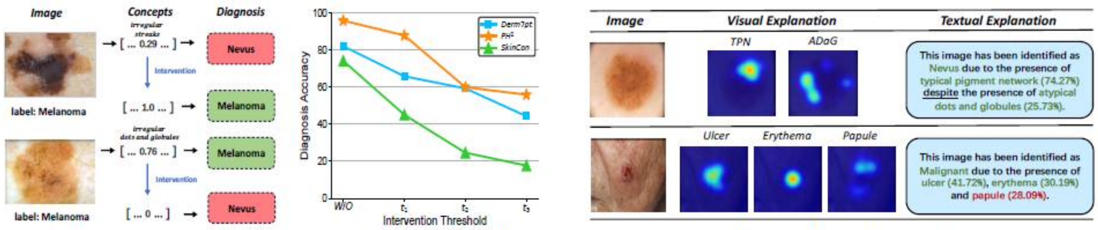
Method learns image and concept semantic correspondences at the image, token, and concept levels

21



Pipeline of proposed framework

21



(a) Intervention examples. (b) Intervention results. (c) Visual and textual explanations.

Model's faithfulness, plausibility and understandability.

- ✓ (a)(b) Test time concept-intervention examples and results.
- ✓ (c) Examples of visual and textual explanations provided by the method given skin images from different datasets.
- ✓ Correct prediction results are marked in green, while red highlights incorrect predictions