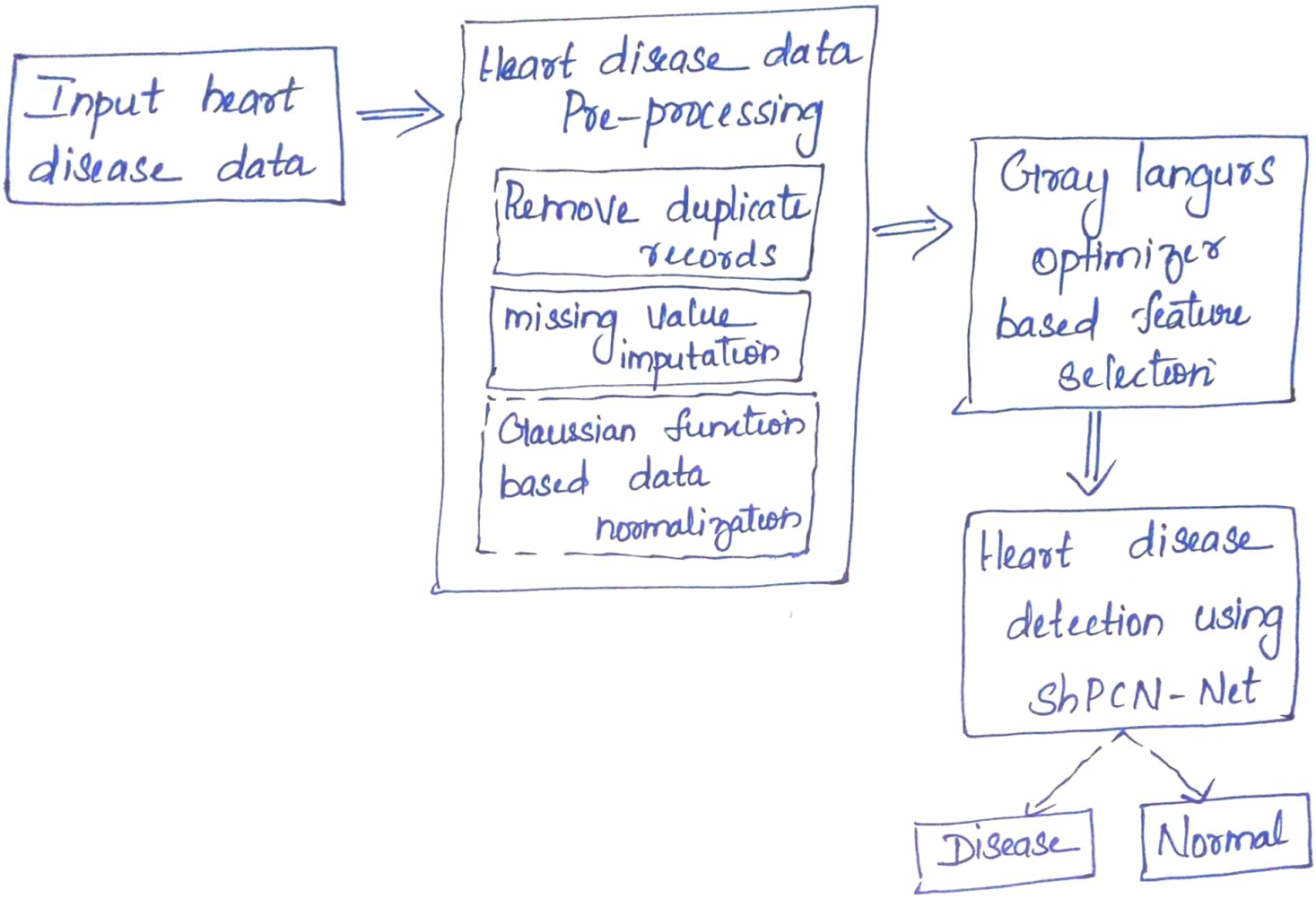
# ShPCN-Net : An Optimization- Based Feature selection and hybrid Learning Framework for Heart disease detection.

## 3. Proposed ShPCN-Net for Heart disease detection




### 3.1. Heart disease data Acquisation

Cleveland heart disease dataset ∅, contains Patient clinical records and diagnostic information for heart disease analysis

$$D = \{d_1, d_2, \ldots\ldots d_i, \ldots d_j\}$$

## 3.2. Preprocessing of Heart disease data

The Heart disease data is preprocessed by handling missing values, removing duplicate records, and ~~encoding~~ data normalization.

### 3.2.1. Remove duplicate records and Handling missing values

⇒ The observational records within the data sets are preprocessed to remove duplicates, missing values that are not relevant.

⇒ Data has been preprocessed for features that contain missing values

⇒ mean imputation is used

$$M_V = \begin{cases} \mu_i & ; \text{if numerical} \\ \text{mode}(i) & ; \text{if categorical} \end{cases}$$

where, $\mu_i$ ⇒ mean of $i$

mode(i) ⇒ most frequent value.

$d_n$ ⇒ [duplicate removal and missing imputation] ⟶ [$M_V$]

Refer: "Advanced intrusion detection in the industrial internet of things using federated learning and LSTM models"

equ. (1)

3.2.2. Gaussian function standardization




$$N_d = \frac{1}{std(M_V)\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{(M_V) - mean(M_V)}{std(M_V)^2}\right)$$

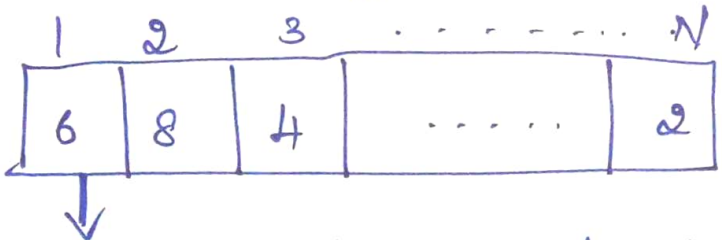
Refer: "Investigating the impact of data normalization methods on predicting electricity consumption in a building using different artifical neural network models"

Section 3.4

3.3. Feature selection using Gray langurs Optimizer.




3.3.1. Solution Encoding




index of selected features.

3.3.2. Fitness function

$$\text{Fitness} = \alpha(1 - \text{Accuracy}) + (1-\alpha)\frac{R}{k}$$

where,

Accuracy → classification accuracy

R → no: of selected features

k → total no: of features.

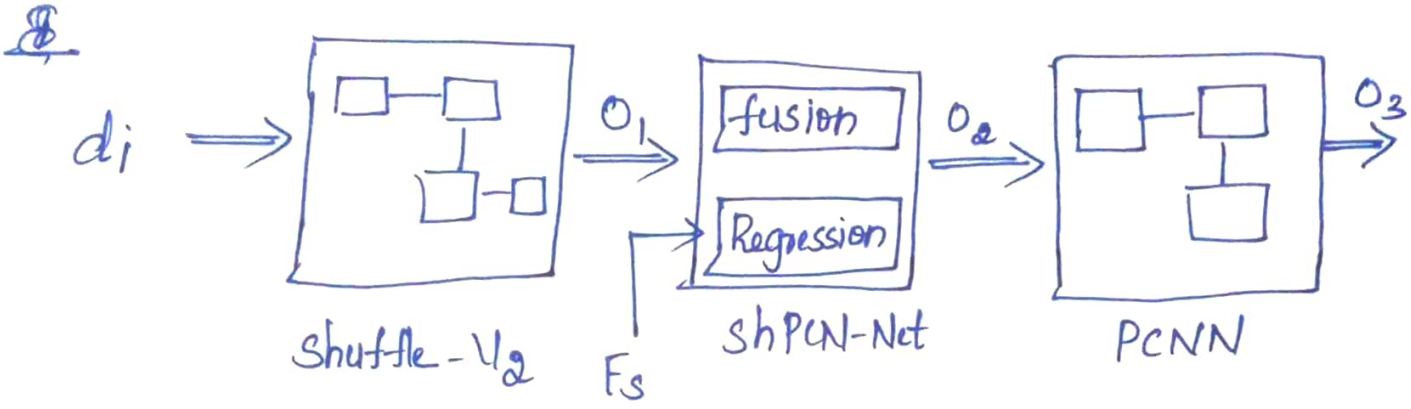
Refer: "Simultaneous SVM parameters and feature selection optimization based on improved slime mould algorithm" equ. (24).

3.3.3. Grey langurs optimizer.

Explain the steps and write the update equation

Refer: "Grey langurs optimizer: a multi-group bio-inspired optimization algorithm".

## 3.4. Heart disease detection using Hybrid DBN.




### 3.4.1. Shuffle Net model.

Refer: "Enhancing missense Variant pathogenicity prediction with missense Net : Integrating structural insights and Shuffle Net - based deep learning Techniques"

Section 4.1 and 4.2, Equation (8)

$$\mathcal{L}_c = \frac{1}{H * W} \sum_{i=1}^{H} \sum_{j=1}^{W} d_i$$

equ (9) $\Rightarrow$ $$s = \sigma\left(g\left(\frac{1}{H*w} \sum_{i=1}^{H} \sum_{j=1}^{W} d_i, w\right)\right)$$

$$s = \sigma\left(W_2\, \delta\left(W_1 \frac{1}{H*W} \sum_{i=1}^{H} \sum_{j=1}^{W} d_i\right)\right)$$

(10) $\Rightarrow$ $$O_1 = \sigma\left(W_2\, \delta\left(W_1 \frac{1}{H_t W_t} \sum_{i=1}^{H} \sum_{j=1}^{W} d_i\right)\right) \cdot d_i$$

3.4.2. Sh PCN-Net

$$y_1 = \sum_{i=1}^{n} F_s * w_i$$

Apply Taylor concept.

$$x(t+1) = x(t) + \frac{x(t)}{1!}$$

$$x(t) = \frac{x(t) - x(t-k)}{k}$$

assume $k=1$, sub $x(t)$ in $x(t+1)$

$$x(t+1) = x(t) + \frac{x(t) - x(t-k)}{1!}$$

$$x(t+1) = 2x(t) - x(t-1)$$

Consider,

$$x(t+1) = O_2, \quad x(t) = y_1, \quad x(t-1) = O_1$$

$$O_2 = 2\left(\sum_{i=1}^{n} F_s * w_i\right) - \sigma\left(w_2 \delta(w_1 \frac{1}{H*W} \sum_{i=1}^{H} \sum_{j=1}^{W} d_i)\right) \cdot d_i$$

3.4.3. PCNN

Refer: "A multiscale parallel convolutional neural network based intelligent human identification

$$O_3 = f_{conv_1}(O_2)$$

$$O_4 = f_{conv_2}(O_2)$$

$$O_5 = f_{conv_3}(O_2)$$

$$O_6 = f_{conv_4}(O_2)$$

$$O_7 = softmax(fc(O_3 || O_4 || O_5 || O_6))$$

## 4. Result and discussion

### 4.1. Tool : python

### 4.2. Metrics : Accuracy, MCC, TPR

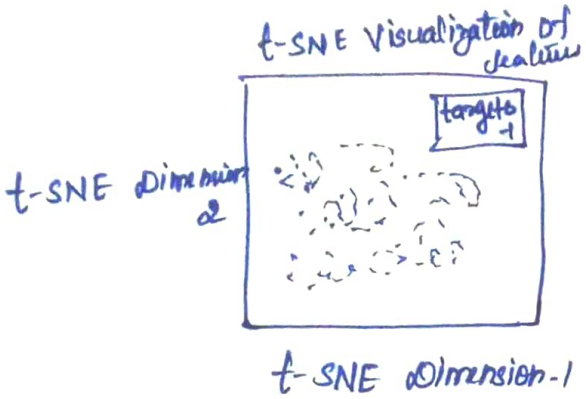
### 4.3. Dataset : cleveland dataset

### 4.4. Parameter for Deep learning and Optimization

| Parameter | range | Parameter | range |
|---|---|---|---|
| hidden layers<br>hidden neurons<br>Activation function<br>Loss function<br>optimizer<br>learning rate<br>Batch size<br>epoch<br>droopout | | Population size<br>maximum iteration<br>control parameter<br>Lower bound<br>upper bound<br>Random number | |

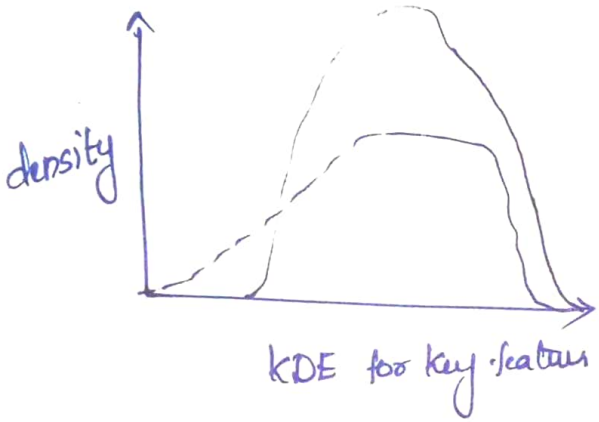
## 4.5. Experimental Results

i) t-SNE




Refer: "predicting cardiovascular risk with hybrid ensemble Learning and explainable AI"

ii) Kernel density Estimation




Refer: "Heart disease prediction using autoencoder and denseNet architecture"

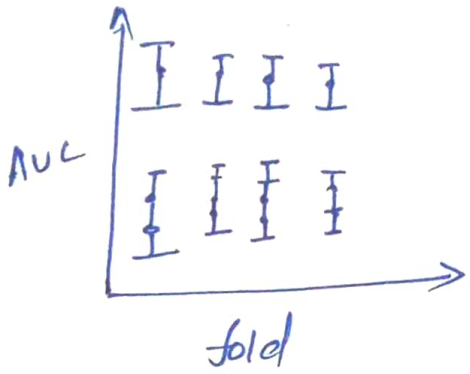
## 4.6. Explainability results using SHap

Refer: "Predicting cardiovascular risk with hybrid ensemble learning and explainable AI"

figure 4

## 4.7. Analysis

1. Fold wise AUC




Refer: "Attention driven graph based machine learning for non-invasive diagnosis of NAFLD"

2. Ablation analysis

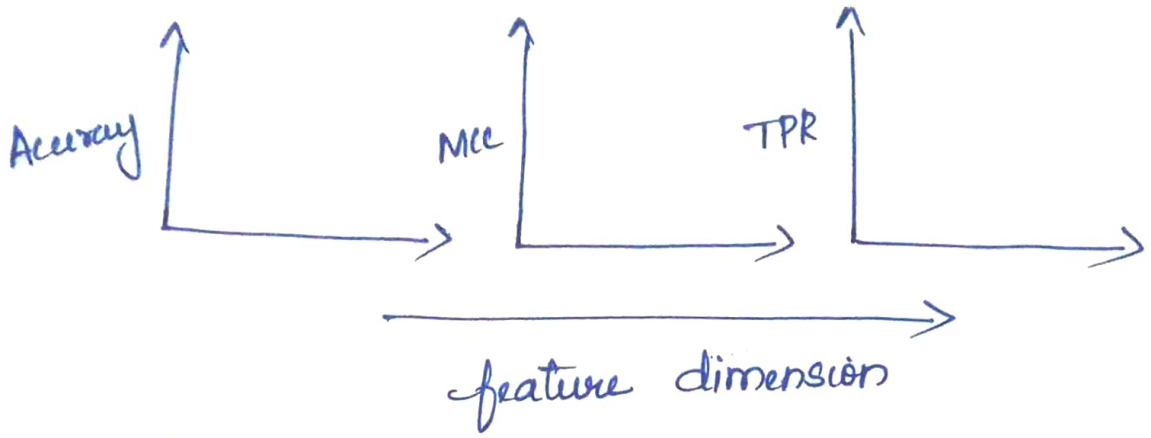



Refer: "Intelligent detection and Early warning of power system cybersecurity Threats based on multimodal large language models"

### methods

1. ShuffleNet
2. PCNN
3. proposed ShPCN-Net

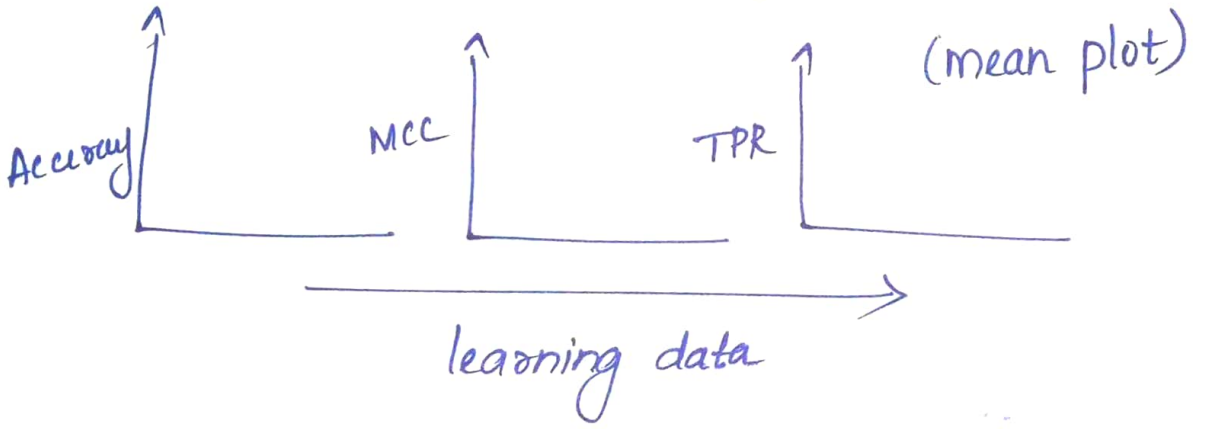
3. Scalability Analysis (violin plot)




methods

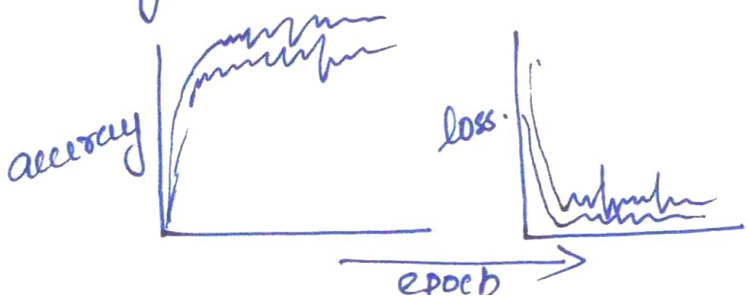
1. HCBiLSTM
2. Dense Autoencoder
3. XGBoost
4. Pooposed shPCN-Net

4. Sensitivity analysis (By Varying 2, 4, 6, 8)




5. confusion matrix

6. Accuracy loss curve




Refer: "Early detection of coronary heart disease using ensemble techniques"

## 4.8. comparative discussion graphs.

Refer: Explainable AI and machine learning: Performance evaluation and explainability of classifiers on educational data mining inspired career counselling"

## 4.9. Wilcoxon rank test

Refer: "Biomodal ECG and PCG cardiovascular disease detection: Exploring the potential and modality contribution" figure - 4.