International Journal of
Mechanical Engineering Research and Technology

ISSN 2454-535X

Email ID: info.ijmert@gmail.com or editor@ijmert.net

ISSN 2454 – 535X
http://www.ijmert.com
Vol.16 Issue 3, 2024

MACHINE LEARNING HEART DISEASE PREDICTION WITH EFFECTIVE FEATURE ENGINEERING

VADHERI HEMANTH KUMAR[1], B RAMA GANESH[2], DUNNA NIKITHA RAO[3]

[1]P.G Scholar, Department of MCA, Sri Venkatesa Perumal College of Engineering & Technology, Puttur, Email: vhemanthkumar26@gmail.com

[2]Professor, Department of CSE, Sri Venkatesa Perumal College of Engineering & Technology, Puttur, Email: ramaganesh34@gmail.com

[3]Assistant Professor, Department of CSE, Sri Padmavathi Mahila Visvavidyalayam Tirupati, Email:rajnikki8195@gmail.com

ABSTRACT: Early diagnosis is vital for heart disease, which influences millions around the world. Principal Component analysis is utilized to find and work on the main qualities in an original element designing technique. The examination utilizes AI to immediately figure Heart Disease condition and make a move. Incorporate undertaking utilizes a Stacking Classifier to coordinate RF, MLP, and LightGBM expectations. This methodology synergistically utilizes model qualities to make a powerful and exact figure with 100 percent accuracy. PCHF-based highlights were utilized to produce the model, and the Stacking Classifier was prepared for front-end arrangement. Flask framework with client validation makes client testing easy and safe, advancing the ML-based heart disease prediction system.

Keywords *– Machine learning, heart failure, cross validations, feature engineering*

1. INTRODUCTION

The heart neglects to siphon sufficient blood to the body in heart failure [1]. Cardiovascular diseases are a worldwide medical problem influencing general wellbeing. Heart failure influences millions worldwide and is hazardous. Ongoing information recommends 26 million individuals experience the ill effects of Heart failure [2]. Two classifications of cardiovascular breakdown exist. A cardiovascular failure or other heart underlying issue. Second, heart-related issues including extreme pulse. Cardiovascular breakdown can cause exhaustion, windedness, and leg and lower leg edema. Drug, way of life adjustments, and medical procedure can treat cardiovascular breakdown. Early ID and treatment of cardiovascular breakdown works on personal satisfaction and endurance [3]. This work fosters a machinelearning model to oversee cardiovascular breakdown to improve patient wellbeing.

ML is critical to clinical conclusion and medical services [4]. ML is utilized in drug advancement, imaging diagnostics, plague forecast, and cardiovascular breakdown expectation. Enormous clinical information might be dissected and anticipated utilizing ML. ML sets aside time and cash, further developing analysis over customary clinical methodologies.[41]

A progressive PCHF highlight designing strategy chooses the main qualities to further develop

International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 535X
http://www.ijmert.com
Vol.16 Issue 3, 2024

execution. Eight dataset qualities with high pertinence values are decided to make PCHF-based ML calculations. We enhanced another list of capabilities to further develop the PCHF system and achieve the best precision scores contrasted with past techniques. Contrast nine strong ML calculations with foresee cardiovascular breakdown. Each ML technique's hyperparameters are tuned to track down the ideal pair, achieving extraordinary exactness. We used k-fold cross-validation to test ML models.

Past investigations have shown that coronary illness is the most deadly human infection. Cardiovascular sicknesses are turning out to be all the more dangerous, compromising medical services frameworks around the world [15], [16]. Generally influenced by this extreme sickness are youngsters [17]. This article [18] covers order models and their utilization in medical care. As per the report, different exploration bunches have effectively assessed information mining approaches in clinical applications. The specialists utilized WEKA and MATLAB to look at useful classifiers. Generally speaking, decision tree, logistic regression, SVM, and different calculations have low precision (52%-67.7%) [19].

As shown in Table 1, past review [11] expanded exactness from 87.27% to 93.13%, which is OK however not ideal. Many methodologies have been utilized to analyze cardiovascular breakdown in patients, including SVM, random forest, decision tree, logistic regression, and naïve bayes classifier Subsequent to looking at discoveries, the decision tree recognized cardiovascular breakdown in a dataset with 93.19% accuracy.[43]

A gathering model for heart diseases finding was built utilizing Cleveland information [20]. The collection model used random forest, gradient boosting, and extreme gradient boosting classifiers to achieve an accuracy of 85.71% [7]. The proposed concentrate on utilized Cleveland information to upgrade coronary illness expectation by include choice, accomplishing 86.60% precision. At last, prior examinations recognized impressive review holes, it is missing to recommend execution exactness. Hence, we thoroughly evaluate the earlier review's presentation investigation here. Results summing up all earlier models' effectiveness illuminate this connected work part. Past exploration show that different models actually conjecture in an unexpected way. Along these lines, dimensionality decrease and element designing further develop information determination and forecast precision [21].

The accuracy score of our proposed study has further developed over the earlier exploration execution score. Legitimate cardiovascular breakdown treatment requires careful capabilities and results. This study utilizes strong ML to do this.

2. LITERATURE REVIEW

More than 26 million individuals overall experience the ill effects of ongoing cardiovascular disease. Cardiovascular disease is responsible for many deaths and more than one million hospitalizations in Europe and North America. Strategies to detect chronic heart failure could prevent, diagnose and avoid hospitalizations and risk conditions, thereby contributing to sustained personal well-being. This examination [1] presents an ML procedure for constant cardiovascular breakdown recognizable proof

International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 535X
http://www.ijmert.com
Vol.16 Issue 3, 2024

utilizing heart sounds. Separating, division, include extraction, and ML are utilized [4, 5, 6, 7, 8, 10]. The technique was assessed with 122 examination volunteers utilizing a leave-one-subject-out approach. The strategy beat a greater part classifier by 15% with 96% accuracy. It recalls 87% of persistent cardiovascular breakdown patients with 87% accuracy. An examination found that strong ML on genuine sounds gathered with a subtle computerized stethoscope can foresee constant cardiovascular breakdown.

Something like 26 million people overall are impacted by heart failure (HF), a developing scourge. HF wellbeing costs are high and will rise extensively with a maturing populace [2]. Indeed, even with forward leaps in drugs and counteraction, mortality and dreariness stay high and personal satisfaction low. In view of HF patients' aetiologies and clinical elements, commonness, occurrence, mortality, and dreariness rates shift topographically [1, 8, 11, 12]. This study covers HF the study of disease transmission worldwide, including commonness, frequency, mortality, and horribleness.

Machine learning (ML)/deep learning (DL) techniques are becoming more commonplace as they demonstrate better performance in a variety of medical applications, from heart failure prediction based on simple cardiac symptoms to computer-aided diagnosis (CADx) based on complex clinical images. In spite of ML/DL's amazing exhibition, there are still questions about its strength in medical care settings (which are customarily troublesome because of the numerous security and protection issues included), particularly after late outcomes showed that ML/DL are powerless against ill-disposed assaults. In this study [4], we examine medical services applications that utilization such comes nearer from a security and protection point of view and their concerns. We additionally propose safe and protection saving ML for medical services applications. At last, we talk about ebb and flow research issues and fascinating future ways.

Coronary heart disease individuals around the world. Heart disease prediction is perhaps of the hardest clinical datum scientific issues. Machine learning(ML) analyze by deciding and foreseeing in light of worldwide medical care information. We likewise saw ML [4, 5, 6, 7, 8, 10] sickness expectation calculations. ML classifiers have been utilized to foresee heart sickness in many examinations [5]. We utilized eleven ML classifiers to find significant factors that improved coronary illness forecast in our article. The forecast model was presented utilizing highlight blends and notable arrangement strategies. Our heart disease prediction model achieved 95% accuracy using gradient boosted trees and multi-layer perceptrons, with random forests being a better predictor of heart disease with 96% accuracy.

Individuals today are engrossed with occupations and different things and disregarding their wellbeing. Because of their bustling lives and wellbeing disregard, more people fall debilitated consistently. Heart disease influences the vast majority. As per WHO figures, near 31% of worldwide fatalities happen from heart disease. In this way, heart disease anticipation is essential for medication. In any case, clinics get such an excess of information that it very well may be challenging to look at. ML [8, 10] can help specialists estimate and oversee information all the more proficiently. In this work [6], we investigated heart sickness, its gamble factors, and ML. We

International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 535X
http://www.ijmert.com
Vol.16 Issue 3, 2024

anticipated heart sickness utilizing ML and analyzed the calculations used for the trial. This venture plans to expect and analyze cardiac disease utilizing ML.[45]

3. METHODOLOGY

i) Proposed Work:

ML can detect heart disease early. We analyze nine ML algorithms: logistic regression, random forest, support vector machine, decision tree, extreme gradient boosting, naive Bayes, k-nearest neighbor, multi-layer perceptron, and gradient boosting. The novel Principal Component Analysis (PCA) includes a design method to select basic attributes to further improve accuracy. In addition, we use a stacking classifier to sum the model expectations of random forest (RF), multi-layer perceptron (MLP) [31], and LightGBM. This technique synergistically utilizes the quality of the models to create a powerful and accurate figure with 100% accuracy. We used PCHF-based highlights to build the model and prepare the stacking classifier for front-end assembly.The Flask framework with client authentication makes client testing easy and secure, advancing ML-based heart disease prediction systems.

ii) System Architecture:

Kaggle gave heart failure dataset to this examination. The assortment incorporates 1025heart failure and healthy patient records. The dataset is organized utilizing information arrangement. Exploratory heart failure information investigation recognizes heart failure factors and patterns. The recommended PCHF technique chooses high-significance highlights in highlight designing. Then the dataset is isolated into train and test. Datasets are handled with nine strong ML strategies. Hyperparameter-put together calibrating is utilized with respect to ML models. The prevalent model predicts heart failure proficiently.

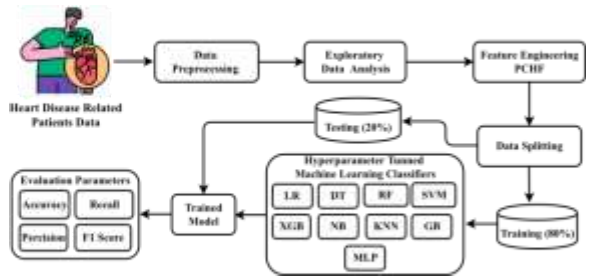

Fig 1 Proposed architecture

iii) Dataset collection:

This examination trains and tests ML calculations for dependable heart disease prediction using a heart disease dataset [39] with demographics, medical history, and physiological parameters.

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
0	52	1	0	125	212	0	1	168	0	1.0	2	2	3	0
1	53	1	0	140	203	1	0	155	1	3.1	0	0	3	0
2	70	1	0	145	174	0	1	125	1	2.6	0	0	3	0
3	61	1	0	148	203	0	1	161	0	0.0	2	1	3	0
4	62	0	0	138	294	1	1	106	0	1.9	1	3	2	0

Fig 2 Heart disease dataset

iv) Data Processing:

Data processing transforms crude information into business-valuable data. Data researchers assemble, sort out, clean, check, investigate, and orchestrate information into charts or papers. Information can be handled physically, precisely, or electronically. Data

International Journal of Mechanical Engineering Research and Technology

ISSN 2454 – 535X
http://www.ijmert.com
Vol.16 Issue 3, 2024

ought to be more significant and decision-production simpler. Organizations might improve tasks and go with basic decisions quicker. PC programming advancement and other mechanized information handling innovations add to this. Big data can be transformed into applicable bits of knowledge for quality administration and independent direction.

v) Feature selection:

Feature selection chooses the most predictable, non-repetitive, and significant elements for model turn of events. As data sets extend in amount and assortment, deliberately bringing down their size is pivotal. The essential reason for highlight determination is to increment prescient model execution and limit processing cost.[47]

One of the vital pieces of feature engineering is picking the main qualities for ML calculations. Feature selection takes out repetitive or pointless attributes and limits the assortment of info factors to those generally vital to the [1, 2] ML model. The significant benefits of pre-choosing attributes instead of letting the ML model choose.

vi) Algorithms:

LR: Logit models are utilized for order and prescient examination. Logistic regression works out the probability of an occasion, like voting, in light of a dataset of free factors [23].

```
# Logistic Regression model
from sklearn.linear_model import LogisticRegression
#from sklearn.pipeline import Pipeline

# instantiate the model
log = LogisticRegression(penalty='l2', fit_intercept=True, random_state = 1, max_iter =100)

log.fit(X_train,y_train)

y_pred = log.predict(X_test)

lr_acc = accuracy_score(y_pred, y_test)
lr_prec = precision_score(y_pred, y_test)
lr_rec = recall_score(y_pred, y_test)
lr_f1 = f1_score(y_pred, y_test)

storeResults('Logistic Regression',lr_acc,lr_prec,lr_rec,lr_f1)
```

Fig 3 Logistic regression

DT: Decision trees are non-parametric managed learning calculations for characterization and relapse. Its tree structure contains a root hub, branches, inner hubs, and leaf hubs.

```
from sklearn.tree import DecisionTreeClassifier

# instantiate the model
tree = DecisionTreeClassifier(criterion='gini',max_depth=[illegible],min_samples_split=[illegible],max_features=None,random_state=[illegible],max_leaf_nodes

tree.fit(X_train, y_train)

y_pred = tree.predict(X_test)

dt_acc = accuracy_score(y_pred, y_test)
dt_prec = precision_score(y_pred, y_test)
dt_rec = recall_score(y_pred, y_test)
dt_f1 = f1_score(y_pred, y_test)

storeResults('Decision Tree',dt_acc,dt_prec,dt_rec,dt_f1)
```

Fig 4 Decision tree

RF: The Leo Breiman and Adele Cutler-protected random forest strategy joins decision tree result to give a solitary result. As a characterization and relapse instrument, its straightforwardness and flexibility have driven its prominence [11].

International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 535X
http://www.ijmert.com
Vol.16 Issue 3, 2024

```python
from sklearn.ensemble import RandomForestClassifier

# instantiate the model
rf = RandomForestClassifier(n_estimators = 100, criterion = 'gini', max_depth=100,  max_features='sqrt',
                            bootstrap = True, random_state = 0, max_samples = None)

rf.fit(X_train, y_train)

y_pred = rf.predict(X_test)

rf_acc = accuracy_score(y_pred, y_test)
rf_prec = precision_score(y_pred, y_test)
rf_rec = recall_score(y_pred, y_test)
rf_f1 = f1_score(y_pred, y_test)

storeResults('Random Forest',rf_acc,rf_prec,rf_rec,rf_f1)
```

Fig 5 Random forest

SVM: Powerful directed calculation SVM performs well on more modest yet confounded datasets. SVMs might be utilized for relapse and arrangement, despite the fact that they perform better in grouping.

```python
from sklearn.svm import SVC

# instantiate the model
svm = SVC(C=1.0,kernel = 'rbf', degree = 3, gamma = 'scale', probability=True, tol = 0.001,cache_size=200,max_iter=-1,random_sta

# fit the model
svm.fit(X_train, y_train)

#predicting the target value from the model for the samples

y_pred = svm.predict(X_test)

svc_acc = accuracy_score(y_pred, y_test)
svc_prec = precision_score(y_pred, y_test)
svc_rec = recall_score(y_pred, y_test)
svc_f1 = f1_score(y_pred, y_test)

storeResults('Support Vector Machine',svc_acc,svc_prec,svc_rec,svc_f1)
```

Fig 6 SVM

KNN: A non-parametric, directed learning classifier, the k-nearest neighbors strategy (KNN) utilizes closeness to characterize or foresee information point gathering.

```python
from sklearn.neighbors import KNeighborsClassifier
#from sklearn.pipeline import Pipeline

# instantiate the model
knn = KNeighborsClassifier(n_neighbors=1, weights='uniform', algorithm='auto',leaf_size=30, p=2, metric='minkowski')

# fit the model
knn.fit(X_train,y_train)

y_pred = knn.predict(X_test)

knn_acc = accuracy_score(y_pred, y_test)
knn_prec = precision_score(y_pred, y_test)
knn_rec = recall_score(y_pred, y_test)
knn_f1 = f1_score(y_pred, y_test)

storeResults('KNN',knn_acc,knn_prec,knn_rec,knn_f1)
```

Fig 7 KNN

MLP: Modern feedforward artificial neural networks, known as multilayer perceptrons (MLPs), are completely connected neurons with a nonlinear enactment capability organized in no less than three layers that can separate nonlinearly detachable info. The first perceptron utilized a Heaviside step capability, not a nonlinear enactment capability like current organizations.[49]

```python
from sklearn.neural_network import MLPClassifier

model = MLPClassifier(hidden_layer_sizes = (5,2),activation='relu',solver = 'lbfgs', alpha = 0.0001,
                      learning_rate = 'constant',random_state=1, max_iter=300, shuffle = True)
# fit the model
model.fit(X_train,y_train)

y_pred = model.predict(X_test)

mlp_acc = accuracy_score(y_pred, y_test)
mlp_prec = precision_score(y_pred, y_test)
mlp_rec = recall_score(y_pred, y_test)
mlp_f1 = f1_score(y_pred, y_test)

storeResults('MLP',mlp_acc,mlp_prec,mlp_rec,mlp_f1)
```

Fig 8 MLP

NB: Naïve Bayes Classifier is a straightforward and powerful methodology for creating quickly prescient ML models. It predicts in view of item probability as a probabilistic classifier.

International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 535X
http://www.ijmert.com
Vol.16 Issue 3, 2024

```
# Naive Bayes Classifier Model
from sklearn.naive_bayes import GaussianNB

# instantiate the model
nb=  GaussianNB(var_smoothing=1e-09)

# fit the model
nb.fit(X_train,y_train)

y_pred = nb.predict(X_test)

nb_acc = accuracy_score(y_pred, y_test)
nb_prec = precision_score(y_pred, y_test)
nb_rec = recall_score(y_pred, y_test)
nb_f1 = f1_score(y_pred, y_test)

storeResults('Naive Bayes',nb_acc,nb_prec,nb_rec,nb_f1)
```

Fig 9 Naïve bayes

XGBoost: XGBoost is a better disseminated gradient boosting toolbox for quick and adaptable ML model preparation. Ensemble learning consolidates powerless model expectations to make a superior forecast.

```
from xgboost import XGBClassifier

# instantiate the model
xgb = XGBClassifier(loss='log_loss', learning_rate=0.1, n_estimators = 100, min_samples_split = 2,
                    min_samples_leaf =1, max_depth = 3, use_label_encoder = False, eval_metric = 'mlogloss')

# fit the model
xgb.fit(X_train,y_train)

y_pred = xgb.predict(X_test)

xgb_acc = accuracy_score(y_pred, y_test)
xgb_prec = precision_score(y_pred, y_test)
xgb_rec = recall_score(y_pred, y_test)
xgb_f1 = f1_score(y_pred, y_test)

storeResults('XGBoost',xgb_acc,xgb_prec,xgb_rec,xgb_f1)
```

Fig 10 XGBoost

Gradient Boosting: Gradient Boosting is a typical ML characterization and relapse approach. Ensemble Learning techniques like supporting train models successively and attempt to address one another. Joining frail students makes serious areas of strength for them.

```
from sklearn.ensemble import GradientBoostingClassifier

# instantiate the model
gbc = GradientBoostingClassifier(learning_rate = 1.0, n_estimators = 20, subsample = 1.0,
                                 criterion = 'friedman_mse', max_depth = 2, random_state = 1)

# fit the model
gbc.fit(X_train,y_train)

y_pred = gbc.predict(X_test)

gb_acc = accuracy_score(y_pred, y_test)
gb_prec = precision_score(y_pred, y_test)
gb_rec = recall_score(y_pred, y_test)
gb_f1 = f1_score(y_pred, y_test)

storeResults('Gradient Boosting',gb_acc,gb_prec,gb_rec,gb_f1)
```

Fig 11 Gradient boosting

Stacking Classifier: A stacking classifier makes a "super" characterization model from many models utilizing ensemble learning. This can boost execution since the joined model can gain from each model's capacities.

```
import lightgbm as lgb
from sklearn.ensemble import StackingClassifier

estimators = [('rf', rf),('dt', tree)]

clf = StackingClassifier(estimators=estimators, final_estimator=lgb.LGBMClassifier(max_depth=-1, random_state=10, silent=True,

clf.fit(X_train, y_train)

y_pred = clf.predict(X_test)

stac_acc = accuracy_score(y_pred, y_test)
stac_prec = precision_score(y_pred, y_test)
stac_rec = recall_score(y_pred, y_test)
stac_f1 = f1_score(y_pred, y_test)

storeResults('Stacking Classifier',stac_acc,stac_prec,stac_rec,stac_f1)
```

Fig 12 Stacking classifier

5. EXPERIMENTAL RESULTS

Accuracy: The model's accuracy is the percentage of true predictions at a grouping position.

$$\text{Accuracy} = TP + TN\ TP + TN + FP + FN.$$

International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 535X
http://www.ijmert.com
Vol.16 Issue 3, 2024

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

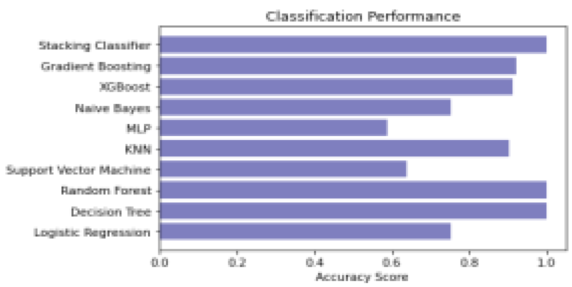

Fig 12 Accuracy graph

Precision: Precision quantifies the percentage of certain events or tests that are well characterized. To attain accuracy, use the formula:

Precision = True positives/ (True positives + False positives) = TP/(TP + FP)

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

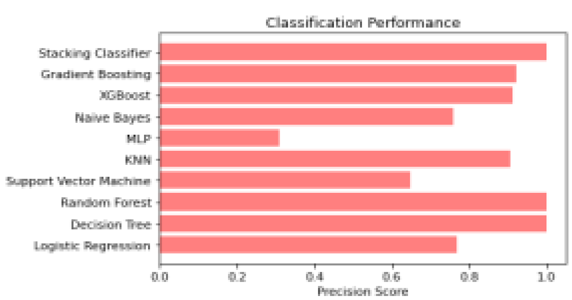

Fig 13 Precision graph

Recall: ML recall measures a model's ability to catch all class occurrences. The model's ability to recognize a certain type of event is measured by the percentage of precisely anticipated positive prospects that turn into real earnings.

$$Recall = \frac{TP}{TP + FN}$$

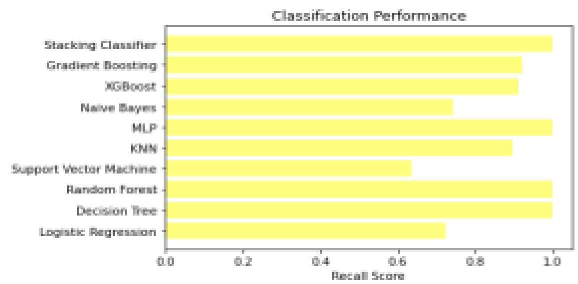

Fig 14 Recall graph

F1-Score: The F1 score captures both false positives and false negatives, making it a harmonized precision and validation technique for unbalanced data sets.

$$\textbf{F1 Score} = \frac{2}{\left(\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}\right)}$$

$$\textbf{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 535X
http://www.ijmert.com
Vol.16 Issue 3, 2024

Fig 15 F1 Score graph

ML Model	Accuracy	F1-score	Recall	Precision
Logistic Regression	0.75	0.816	0.724	0.767
Decision Tree	1.000	1.000	1.000	1.000
Random Forest	1.000	1.000	1.000	1.000
SVM	0.639	0.660	0.636	0.648
KNN	0.902	0.193	0.895	0.904
MLP	0.590	0.184	1.000	0.311
Naive Bayes	0.751	0.777	0.741	0.758
XG Boosting	0.912	0.913	0.913	0.913
Gradient Boosting	0.922	9.22	0.922	0.922
Staking Classifier	1.000	1.000	1.000	1.000

Fig 16 Performance Evaluation

Fig 17 Home page

Fig 18 Signup page

Fig 19 Signin page

Fig 20 Upload input values to predict result

Fig 21 Predict result as you have no heart disease, based on the input provide

6. CONCLUSION

This study proposes ML prediction of heart failure [22]. The model is created using 1,025 patient records. Another PCHF technique to create an inclusion design selects the eight most important attributes to further improve the execution. This study uses logistic regression, random forest, support vector machine,

International Journal of Mechanical Engineering Research and Technology

ISSN 2454 – 535X
http://www.ijmert.com
Vol.16 Issue 3, 2024

decision tree, extreme gradient boosting, simple basis, k-nearest neighbor, multi-layer perceptron, and gradient boosting. With a running time calculation of 0.005, the proposed DT method was 100% accurate. Each learning model presentation is checked using a 10-level mutual acceptance test. Our approach to heart failure detection outperforms state-of-the-art tests and is generalizable.

7. FUTURE SCOPE

Our methodologies can set an exhibition benchmark for heart disease prediction, guiding future examination. To further develop arrangement models, future examination could refine highlight the executives. Our methodology could likewise further develop disease prediction and analysis using ML techniques [1, 2, 4, 10].

REFERENCES

[1] M. Gjoreski, M. Simjanoska, A. Gradišek, A. Peterlin, M. Gams, and G. Poglajen, ‘‘Chronic heart failure detection from heart sounds using a stack of machine-learning classifiers,’’ in Proc. Int. Conf. Intell. Environments (IE), Aug. 2017, pp. 14–19.

[2] G. Savarese and L. H. Lund, ‘‘Global public health burden of heart failure,’’ Cardiac Failure Rev., vol. 3, no. 1, p. 7, 2017.

[3] E. J. Benjamin et al., ‘‘Heart disease and stroke statistics—2019 update: A report from the American heart association,’’ Circulation, vol. 139, no. 10, pp. e56–e528, 2019.

[4] A. Qayyum, J. Qadir, M. Bilal, and A. Al-Fuqaha, ‘‘Secure and robust machine learning for healthcare: A survey,’’ IEEE Rev. Biomed. Eng., vol. 14, pp. 156–180, 2021.

[5] C. A. U. Hassan, J. Iqbal, R. Irfan, S. Hussain, A. D. Algarni, S. S. H. Bukhari, N. Alturki, and S. S. Ullah, ‘‘Effectively predicting the presence of coronary heart disease using machine learning classifiers,’’ Sensors, vol. 22, no. 19, p. 7227, Sep. 2022.

[6] R. Katarya and S. K. Meena, ‘‘Machine learning techniques for heart disease prediction: A comparative study and analysis,’’ Health Technol., vol. 11, no. 1, pp. 87–97, Jan. 2021.

[7] P. Rani, R. Kumar, N. M. O. S. Ahmed, and A. Jain, ‘‘A decision support system for heart disease prediction based upon machine learning,’’ J. Reliable Intell. Environments, vol. 7, no. 3, pp. 263–275, Sep. 2021.

[8] N. S. Mansur Huang, Z. Ibrahim, and N. Mat Diah, ‘‘Machine learning techniques for early heart failure prediction,’’ Malaysian J. Comput. (MJoC), vol. 6, no. 2, pp. 872–884, 2021.

[9] T. Amarbayasgalan, V. Pham, N. Theera-Umpon, Y. Piao, and K. H. Ryu, ‘‘An efficient prediction method for coronary heart disease risk based on two deep neural networks trained on well-ordered training datasets,’’ IEEE Access, vol. 9, pp. 135210–135223, 2021.

[10] R. Bharti, A. Khamparia, M. Shabaz, G. Dhiman, S. Pande, and P. Singh, ‘‘Prediction of heart disease using a combination of machine learning and deep learning,’’ Comput. Intell. Neurosci., vol. 2021, pp. 1–11, Jul. 2021.

International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 535X
http://www.ijmert.com
Vol.16 Issue 3, 2024

[11] F. S. Alotaibi, ‘‘Implementation of machine learning model to predict heart failure disease,’’ Int. J. Adv. Comput. Sci. Appl., vol. 10, no. 6, pp. 1–8, 2019.

[12] D. K. Plati, E. E. Tripoliti, A. Bechlioulis, A. Rammos, I. Dimou, L. Lakkas, C. Watson, K. McDonald, M. Ledwidge, R. Pharithi, J. Gallagher, L. K. Michalis, Y. Goletsis, K. K. Naka, and D. I. Fotiadis, ‘‘A machine learning approach for chronic heart failure diagnosis,’’ Diagnostics, vol. 11, no. 10, p. 1863, Oct. 2021.

[13] A. Saboor, M. Usman, S. Ali, A. Samad, M. F. Abrar, and N. Ullah, ‘‘A method for improving prediction of human heart disease using machine learning algorithms,’’ Mobile Inf. Syst., vol. 2022, pp. 1–9, Mar. 2022.

[14] S. Sarah, M. K. Gourisaria, S. Khare, and H. Das, ‘‘Heart disease prediction using core machine learning techniques—A comparative study,’’ in Advances in Data and Information Sciences. Springer, 2022, pp. 247–260.

[15] C. Trevisan, G. Sergi, and S. Maggi, ‘‘Gender differences in brain-heart connection,’’ Brain and Heart Dynamics. 2020, pp. 937–951.

[16] M. S. Oh and M. H. Jeong, ‘‘Sex differences in cardiovascular disease risk factors among Korean adults,’’ Korean J. Med., vol. 95, no. 4, pp. 266–275, Aug. 2020.

[17] D. C. Yadav and S. Pal, ‘‘Prediction of heart disease using feature selection and random forest ensemble method,’’ Int. J. Pharmaceutical Res., vol. 12, no. 4, pp. 56–66, 2020.

[18] D. Tomar and S. Agarwal, ‘‘A survey on data mining approaches for healthcare,’’ Int. J. Bio-Sci. Bio-Technol., vol. 5, no. 5, pp. 241–266, Oct. 2013.

[19] S. Ekiz and P. Erdogmus, ‘‘Comparative study of heart disease classification,’’ in Proc. Electr. Electron., Comput. Sci., Biomed. Eng. Meeting (EBBT), Apr. 2017, pp. 1–4.

[20] B. A. Tama, S. Im, and S. Lee, ‘‘Improving an intelligent detection system for coronary heart disease using a two-tier classifier ensemble,’’ BioMed Res. Int., vol. 2020, pp. 1–10, Apr. 2020.

[21] V. Ramalingam, A. Dandapath, and M. K. Raja, ‘‘Heart disease prediction using machine learning techniques: A survey,’’ Int. J. Eng. Technol., vol. 7, no. 2, pp. 684–687, 2018.

[22] Heart Disease Dataset|Kaggle, DAVID LAPP, Atlanta, Georgia, 1988.

[23] K. Shah, H. Patel, D. Sanghvi, and M. Shah, ‘‘A comparative analysis of logistic regression, random forest and KNN models for the text classification,’’ Augmented Human Res., vol. 5, no. 1, pp. 1–16, Dec. 2020.

[24] M. M. Ali, B. K. Paul, K. Ahmed, F. M. Bui, J. M. W. Quinn, and M. A. Moni, ‘‘Heart disease prediction using supervised machine learning algorithms: Performance analysis and comparison,’’ Comput. Biol. Med., vol. 136, Sep. 2021, Art. no. 104672.

[25] A. Raza, K. Munir, M. Almutairi, F. Younas, and M. M. S. Fareed, ‘‘Predicting employee attrition using machine learning approaches,’’ Appl. Sci., vol. 12, no. 13, p. 6424, Jun. 2022.

International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 535X
http://www.ijmert.com
Vol.16 Issue 3, 2024

[26] A. Raza, H. U. R. Siddiqui, K. Munir, M. Almutairi, F. Rustam, and I. Ashraf, ''Ensemble learning-based feature engineering to analyze maternal health during pregnancy and health risk prediction,'' PLoS ONE, vol. 17, no. 11, Nov. 2022, Art. no. e0276525.

[27] S. Shabani, S. Samadianfard, M. T. Sattari, A. Mosavi, S. Shamshirband, T. Kmet, and A. R. Várkonyi-Kóczy, ''Modeling pan evaporation using Gaussian process regression K-nearest neighbors random forest and support vector machines; comparative analysis,'' Atmosphere, vol. 11, no. 1, p. 66, Jan. 2020.

[28] N. L. Fitriyani, M. Syafrudin, G. Alfian, and J. Rhee, ''HDPM: An effective heart disease prediction model for a clinical decision support system,'' IEEE Access, vol. 8, pp. 133034–133050, 2020.

[29] S. Bashir, Z. S. Khan, F. H. Khan, A. Anjum, and K. Bashir, ''Improving heart disease prediction using feature selection approaches,'' in Proc. 16th Int. Bhurban Conf. Appl. Sci. Technol. (IBCAST), Jan. 2019, pp. 619–623.

[30] D. D. Rufo, T. G. Debelee, A. Ibenthal, and W. G. Negera, ''Diagnosis of diabetes mellitus using gradient boosting machine (LightGBM),'' Diagnostics, vol. 11, no. 9, p. 1714, Sep. 2021.

[31] R. Pahuja and A. Kumar, ''Sound-spectrogram based automatic bird species recognition using MLP classifier,'' Appl. Acoust., vol. 180, Sep. 2021, Art. no. 108077.

[32] U. Azmat, Y. Y. Ghadi, T. A. Shloul, S. A. Alsuhibany, A. Jalal, and J. Park, ''Smartphone sensor-based human locomotion surveillance system using multilayer perceptron,'' Appl. Sci., vol. 12, no. 5, p. 2550, Feb. 2022.

[33] J. Isabona, A. L. Imoize, and Y. Kim, ''Machine learning-based boosted regression ensemble combined with hyperparameter tuning for optimal adaptive learning,'' Sensors, vol. 22, no. 10, p. 3776, May 2022.

[34] A. A. Ali, H. S. Hassan, and E. M. Anwar, ''Heart diseases diagnosis based on a novel convolution neural network and gate recurrent unit technique,'' in Proc. 12th Int. Conf. Electr. Eng. (ICEENG), Jul. 2020, pp. 145–150.

[35] O. E. Taylor, P. S. Ezekiel, and F. B. D. Okuchaba, ''A model to detect heart disease using machine learning algorithm,'' Int. J. Comput. Sci. Eng., vol. 7, no. 11, pp. 1–5, Nov. 2019.

[36] D. K. Chohan and D. C. Dobhal, ''A comparison based study of supervised machine learning algorithms for prediction of heart disease,'' in Proc. Int. Conf. Comput. Intell. Sustain. Eng. Solutions (CISES), May 2022, pp. 372–375.

[37] K. Sanjar, O. Bekhzod, J. Kim, A. Paul, and J. Kim, ''Missing data imputation for geolocation-based price prediction using KNN–MCF method,'' ISPRS Int. J. Geo-Inf., vol. 9, no. 4, p. 227, Apr. 2020.

[38] B. Olimov, B. Subramanian, R. A. A. Ugli, J.-S. Kim, and J. Kim, ''Consecutive multiscale feature learning-based image classification model,'' Sci. Rep., vol. 13, no. 1, p. 3595, Mar. 2023.

[39] B. Olimov, S. Karshiev, E. Jang, S. Din, A. Paul, and J. Kim, ''Weight initialization based-rectified linear unit activation function to improve the

ISSN 2454 – 535X
http://www.ijmert.com
Vol.16 Issue 3, 2024

performance of a convolutional neural network model,’’ Concurrency Comput., Pract. Exp., vol. 33, no. 22, p. e6143, Nov. 2021.

[40] G.Viswanath, “Hybrid encryption framework for securing big data storage in multi-cloud environment”, Evolutionary intelligence, vol.14, 2021, pp.691-698.

[41] Viswanath Gudditi, “Adaptive Light Weight Encryption Algorithm for Securing Multi-Cloud Storage”, Turkish Journal of Computer and Mathematics Education (TURCOMAT), vol.12, 2021, pp.545-552.

[42] Viswanath Gudditi, “A Smart Recommendation System for Medicine using Intelligent NLP Techniques”, 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS), 2022, pp.1081-1084.

[43] G.Viswanath, “Enhancing power unbiased cooperative media access control protocol in manets”, International Journal of Engineering Inventions, 2014, vol.4, pp.8-12.

[44] Viswanath G, “A Hybrid Particle Swarm Optimization and C4.5 for Network Intrusion Detection and Prevention System”, 2024, International Journal of Computing, DOI: https://doi.org/10.47839/ijc.23.1.3442, vol.23, 2024, pp.109-115.

[45] G.Viswanath, “A Real Time online Food Ording application based DJANGO Restfull Framework”, Juni Khyat, vol.13, 2023, pp.154-162.

[46] Gudditi Viswanath, “Distributed Utility-Based Energy Efficient Cooperative Medium Access Control in MANETS”, 2014, International Journal of Engineering Inventions, vol.4, pp.08-12.

[47] G.Viswanath,“ A Real-Time Video Based Vehicle Classification, Detection And Counting System”, 2023, Industrial Engineering Journal, vol.52, pp.474-480.

[48] G.Viswanath, “A Real- Time Case Scenario Based On Url Phishing Detection Through Login Urls ”, 2023, Material Science Technology, vol.22, pp.103-108.

[49] Manmohan Singh,Susheel Kumar Tiwari, G. Swapna, Kirti Verma, Vikas Prasad, Vinod Patidar, Dharmendra Sharma and Hemant Mewada, “A Drug-Target Interaction Prediction Based on Supervised Probabilistic Classification” published in Journal of Computer Science, Available at: https://pdfs.semanticscholar.org/69ac/f07f2e756b79181e4f1e75f9e0f275a56b8e.pdf