A Predictive Analysis of Heart Rates Using Machine Learning

Dr.U.Sivaji
Information Technology
Institute of Aeronautical Engineering
Hyderabad, India
usivaji@iare.ac.in

Panchamahalkar Maithili
Information Technology
Institute of Aeronautical Engineering
Hyderabad, India
20951a1239@iare.ac.in

Ashrith Reddy
Information Technology
Institute of Aeronautical Engineering
Hyderabad, India
20951a1201@iare.ac.in

Thumu Sruthi
Information Technology
Institute of Aeronautical Engineering
Hyderabad, India
20951a12a7@iare.ac.in

ABSTRACT: Heart disease, a chronic illness that influences millions around the world, underscores the need of early distinguishing proof. An extraordinary element designing methodology, utilizing Principal Component Analysis, is given to find and work on the main qualities. The's undertaking will probably utilize ML to gauge the wellbeing condition of individuals with heart disease and take important intercessions. In this review, a group approach, particularly a Stacking Classifier, is utilized to coordinate the forecasts of Random Forest (RF), Multilayer Perceptron (MLP), and LightGBM models. This methodology utilizes the qualities of various models to create an exceptionally versatile and exact last gauge, with an astonishing 100 percent accuracy. The chose highlights in light of Principal Component Heart Failure (PCHF) were utilized to foster the model, and the Stacking Classifier was prepared to be sent in the front end. The blend of the Flask framework with client validation gives a powerful and safe stage for client testing, expanding the availability and convenience of our 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 illnesses are a worldwide medical problem influencing general wellbeing. Cardiovascular breakdown influences millions worldwide and is risky. Ongoing information proposes 26 million individuals experience the ill effects of cardiovascular breakdown [2]. Two classifications of cardiovascular breakdown exist. A cardiovascular failure or other heart primary issue. Second, heart-related issues including extreme pulse. Cardiovascular breakdown can cause exhaustion, windedness, and leg and lower leg edema. Medicine, way of life adjustments, and medical procedure can treat cardiovascular breakdown. Early distinguishing proof 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 vital to clinical finding and medical services [4]. ML is utilized in drug advancement, imaging diagnostics, scourge forecast, and cardiovascular breakdown expectation. Enormous clinical information might be broke down and anticipated utilizing ML. ML sets aside time and cash, further developing analysis over conventional clinical methodologies.

A progressive PCHF highlight designing strategy chooses the main qualities to further develop execution. Eight dataset attributes 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 strategies. Contrast nine strong AI calculations with foresee cardiovascular breakdown. Each ML technique's hyperparameters are tuned to track down the ideal pair, achieving extraordinary precision. We used k-fold cross-validation to test ML models.

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

As exhibited in Table 1, past review [11] expanded accuracy from 87.27% to 93.13%, which is adequate yet 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 In the wake of looking at discoveries, the decision tree distinguished cardiovascular breakdown in a dataset with 93.19% precision.

A group model for coronary illness conclusion was built utilizing Cleveland information [20]. The group models utilized random forest, gradient boosting, and extreme gradient boosting classifiers to accomplish 85.71% accuracy [7]. The proposed concentrate on utilized Cleveland information to upgrade coronary illness expectation by highlight determination, accomplishing 86.60% accuracy. At long last, prior examinations recognized impressive review holes, it is missing to propose execution accuracy. Thusly, we extensively evaluate the earlier review's presentation examination here. Results summing up all earlier models' effectiveness illuminate this connected work part. Past exploration show that different models actually figure in an unexpected way. In this manner, dimensionality decrease and element designing further develop information determination and expectation accuracy[21].

The accuracy score of our proposed study has further developed over the earlier exploration execution score. Appropriate cardiovascular breakdown treatment requires precise 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 persistent cardiovascular breakdown.

It kills numerous cardiovascular illness patients and causes more than 1 million hospitalizations in Europe and North America. Persistent cardiovascular breakdown recognition strategies can forestall, analyze, and forestall hospitalizations and hazardous conditions, working on understanding personal satisfaction. This examination [1] presents an ML procedure for constant cardiovascular breakdown recognizable proof utilizing heart sounds. Sifting, division, include extraction, and ML are utilized [4, 5, 6, 7, 8, 10]. The procedure was assessed with 122 exploration volunteers utilizing a leave-one-subject-out approach. The procedure beat a greater part classifier by 15% with 96% accuracy. It reviews 87% of constant cardiovascular breakdown patients with 87% accuracy. An exploration found that strong ML on genuine sounds gathered with an unnoticeable computerized stethoscope can foresee persistent cardiovascular breakdown.

Something like 26 million people overall are impacted by heart failure (HF), a developing pandemic. HF wellbeing costs are high and will rise extensively with a maturing populace [2]. Indeed, even with forward leaps in meds and counteraction, mortality and bleakness stay high and personal satisfaction low. In light of HF patients' aetiologies and clinical highlights, commonness, frequency, mortality, and dismalness rates fluctuate geologically [1, 8, 11, 12]. This study covers HF the study of disease transmission worldwide, including predominance, rate, mortality, and grimness.

Since machine learning (ML)/deep learning (DL) procedures perform better for an assortment of medical services applications, from heart failure expectation from one-layered heart signs to computer-aided diagnosis (CADx) utilizing multi-faceted clinical pictures, they have become broadly utilized. Notwithstanding ML/DL's noteworthy presentation, there are still questions about its heartiness in medical care settings (which are customarily troublesome because of the numerous security and protection issues included), particularly after ongoing outcomes showed that ML/DL are powerless against antagonistic assaults. In this study [4], we examine medical care applications that utilization such comes nearer from a security and protection point of view and their concerns. We likewise propose safe and protection saving ML for medical services applications. At long last, we talk about momentum research issues and intriguing future ways.

Coronary heart disease kills individuals around the world. Heart disease prediction is quite possibly of the hardest clinical datum scientific issues. ML analyze by simply deciding and anticipating in light of worldwide medical services information. We additionally saw ML [4, 5, 6, 7, 8, 10] ailment expectation calculations. ML classifiers have been utilized to anticipate cardiovascular sickness in many examinations [5]. We utilized eleven ML classifiers to find essential factors that improved coronary illness forecast in our article. The forecast model was presented utilizing highlight mixes and notable arrangement strategies. Our coronary illness expectation model had 95% accuracy with inclination supported trees and multi-layer perceptron. With 96% accuracy, the Random Forest predicts coronary illness better.

Individuals today are distracted with occupations and different things and dismissing their wellbeing. Because of their bustling lives and wellbeing disregard, more people fall debilitated consistently.

Coronary illness influences the vast majority. As per WHO figures, near 31% of worldwide fatalities happen from coronary illness. Accordingly, coronary illness anticipation is critical for medication. In any case, emergency clinics get such an excess of information that it very well may be challenging to look at. ML[8, 10] can help specialists figure and oversee information all the more proficiently. In this work [6], we investigated cardiovascular illness, its gamble factors, and ML. We anticipated cardiovascular infection utilizing ML and thought about the calculations used for the analysis. This task intends to expect and investigate heart sickness utilizing ML.

3. METHODOLOGY

i) Proposed Work:

Machine learning can detect heart disease early. We examine nine machine learning algorithms: logistic regression, random forest, support vector machine, decision tree, extreme gradient boosting, naive bayes, k-nearest neighbors, multilayer perceptron, and gradient boosting. An novel Principal Component Analysis (PCA) feature engineering approach selects critical characteristics to improve accuracy. Additionally, a Stacking Classifier is used to aggregate Random Forest (RF), Multilayer Perceptron (MLP) [31], and LightGBM model predictions. This strategy synergistically uses model strengths to make a robust and accurate forecast with 100% accuracy. PCHF-based features were used to generate the model, and the Stacking Classifier was trained for front-end deployment. Flask framework with user authentication makes user testing easy and secure, improving our machine learning-based heart disease prediction system.

ii) System Architecture:

Kaggle provided heart failure dataset for this investigation. The collection includes 1025 heart failure and healthy patient records. The dataset is formatted using data preparation. Exploratory heart failure data analysis helps identify heart failure factors and trends. The suggested PCHF method selects high-importance features in feature engineering. Then the dataset is separated into train and test. Datasets are processed with nine powerful machine-learning methods. Hyperparameter-based fine tuning is used on machine learning models. The superior model predicts heart failure efficiently.

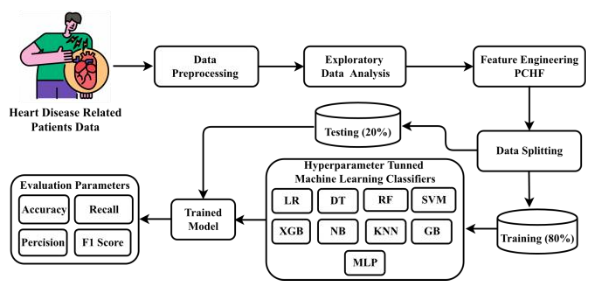

Fig 1 Proposed architecture

iii) Dataset collection:

This research trains and tests machine learning algorithms for reliable 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 turns raw data into business-useful information. Data scientists gather, organize, clean, verify, analyze, and arrange data into graphs or papers. Data can be processed manually, mechanically, or electronically. Information should be more valuable and decision-making easier. Businesses may enhance operations and make critical choices faster. Computer software development and other automated data processing technologies contribute to this. Big data can be turned into relevant insights for quality management and decision-making.

v) Feature selection:

Feature selection chooses the most predictable, non-excess, and applicable elements for model turn of events. As data sets grow in amount and assortment, deliberately bringing down their size is significant. The essential reason for include choice is to increment prescient model execution and limit processing cost.

One of the critical pieces of feature selection is picking the main attributes for ML calculations. To diminish input factors, feature selection techniques wipe out copy or superfluous elements and limit the assortment to those generally critical to the ML model [1, 2]. Rather than permitting the ML model pick the main qualities, highlight determination ahead of time enjoys a few benefits.

vi) Algorithms:

LR: The logit model is utilized for order and prescient examination. Logistic regression utilizes a dataset of free factors to gauge the probability of an event, for example, voting [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 directed learning calculations for arrangement and relapse. Its tree structure contains a root hub, branches, inside hubs, and leaf hubs.

```
from sklearn.tree import DecisionTreeClassifier

# instantiate the model
tree = DecisionTreeClassifier(criterion='gini',max_depth=300,min_samples_split=2,max_features=None,random_state=0,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-licensed random forest strategy consolidates decision tree result to give a solitary result. As an order and relapse instrument, its straightforwardness and flexibility have driven its notoriety [11].

```python
from sklearn.ensemble import RandomForestClassifier

# instantiate the model
rf = RandomForestClassifier(n_estimators = 300, criterion = 'gini', max_depth=300,  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: Strong supervised calculation SVM performs well on more modest yet confounded datasets. SVMs might be utilized for relapse and order, in spite of the fact that they perform better in arrangement.

```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, supervised learning classifier, the k-nearest neighbors method (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=3, 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 distinguishable information. The first perceptron utilized a Heaviside step capability, not a nonlinear initiation capability like current organizations.

```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 basic and successful methodology for producing quickly prescient ML models. It predicts in light of article probability as a probabilistic classifier.

```python
# 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 superior disseminated gradient boosting toolkit for quick and versatile ML model preparation. Ensemble learning consolidates feeble model forecasts to make a superior expectation.

```python
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 order and relapse approach. Ensemble Learning techniques like boosting train models sequentially and attempt to address one another. Consolidating feeble students makes areas of strength for them.

```python
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" grouping model from many models utilizing ensemble learning. This can support execution since the consolidated model can gain from each model's capacities.

```python
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=314, 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: Accuracy is characterized as the extent of right forecasts in a grouping position, which estimates a model's general accuracy.

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

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

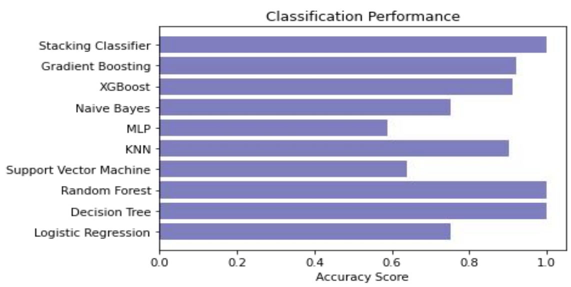

Fig 12 Accuracy graph

Precision: Precision estimates the extent of precisely characterized cases or tests among those classified as certain. Hence, the precision can be determined utilizing the accompanying formula:

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

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

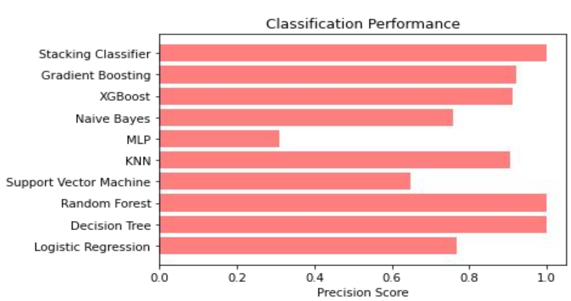

Fig 13 Precision graph

Recall: Recall is an ML metric that evaluates a model's capacity to perceive all occasions of a given class. It is the proportion of accurately anticipated positive perceptions to add up to real up-sides, which gives data on a model's fulfillment in gathering instances of a particular class.

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

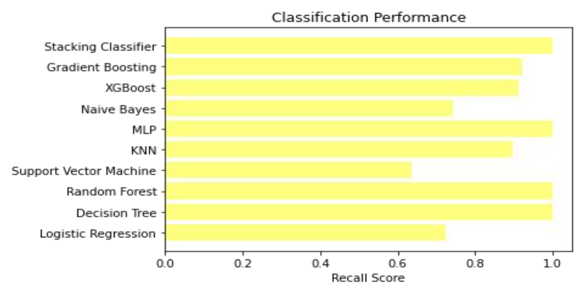

Fig 14 Recall graph

F1-Score: The F1 score is a symphonic medium of precision and validation, and is legitimate for unbalanced datasets because it represents a useful measure that captures both false positives and false negatives.

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

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

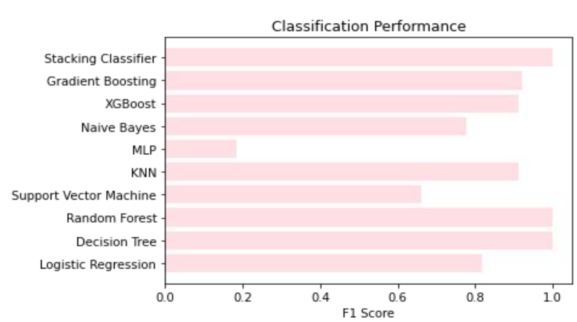

Fig 15 F1 Score graph

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

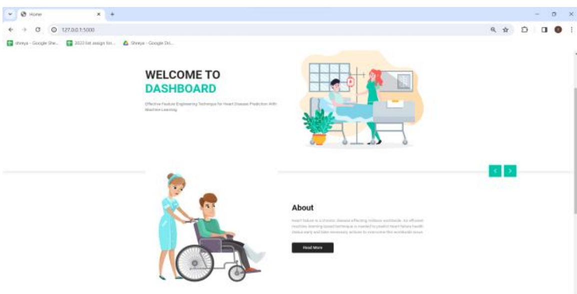

Fig 17 Home page

Fig 18 Signup page

Fig 19 Signin page

Fig 20 Upload input values to predict result

Result: You have no Heart Disease, based on the input provide!

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

6. CONCLUSION

This work proposes machine learning heart failure prediction [22]. Models are built using 1025 patient records. A new PCHF feature engineering method picks the eight most important characteristics to improve performance. In comparison, logistic regression, random forest, support vector machine, decision tree, extreme gradient boosting, naive base, k-nearest neighbors, multilayer perceptron, and gradient boosting are used. With 0.005 runtime computations, the proposed DT technique was 100% accurate. Each learning model's performance is validated using 10-fold cross-validation. Our heart failure detection approach outperformed state-of-the-art investigations and is generalized.

7. FUTURE SCOPE

The findings obtained using our suggested methodologies may be used to set a performance standard for cardiac disease prediction, serving as a benchmark for future study in this area. Subsequent research might focus on improving the feature management method to improve the efficacy of classification models. Furthermore, our technique has the potential to be used in a variety of medical areas to improve illness prediction and identification through the use of machine learning algorithms [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.

[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.

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