Cp Based Heart Attack Detection Using Stacking Classifier

Mr. Mohit Patel[1], Dr. Dushyantsinh B. Rathod[2]
PhD Scholar1 ,Professor & HOD2
,Computer Engineering Department1,2
Swarnim Startup & Innovation University,Kalol,Gujarat1,
Ahmedabad Institute of Technology,Ahmedabad,Gujarat2

Abstract-Heart Disease is one of the foremost common diseases nowadays, and for people that provide health care, it's very necessary to figure out with them to require care of their patients' health and save their life. during this paper, different classifiers were analyzed by performance comparison to classify the guts Disease dataset to classify it correctly and or to Predict heart condition cases with minimal attributes. Large amounts of knowledge that contain some secret information were collected by the healthcare industries. This data collection is beneficial for creating effective decisions. During this case, a Heart Disease Prediction System (HDPS) is developed using Logistic Regression, K Nearest Neighbor, Decision Tree, Random Forest Classifier, and Support Vector Machine algorithms to predict the guts disease risk level.
Keywords: Heart Disease, Support Vector Machine, Accuracy, Decision Tree.

I. Introduction

The most vital and essential part (organ) of the mortal body is the Heart. There are numerous conditions that are linked to the heart so the analysis of vaticination of the heart must be accurate. To resolve this, virtual study about this field is obligatory. Typically, these conditions are prognosticated at the end stage and this is the main reason for the death of heart cases due to insufficiency of correctness because of this there's a need to identify complete algorithms for conditions vaticination. One of the complete able and effective technologies in Machine Literacy that have established on specific training and testing with the support of python and python libraries. The system acquires training directly from data and skill, grounded on this training, testing should be done on colorful types of need as per needful algorithms. For Testing and Training, Machine literacy can be used in an effective fashion. Disease-related to the guts, also appertained to as complaint (CVD), discusses colorful conditions that affect the guts not just the complaint. Heart complaints of the coronary highways, cardiomyopathy, and cardiovascular health issues are certain services where the blood is pumped and its rotation is formed throughout the body. The opinion is an important task that has to be performed efficiently. This is substantially done under a croaker's guidance. This causes wrong results & inordinate medical costs of treatments handed to cases. So, we conclude that an automatic opinion and vaticination system would prove extremely favorable.

Nowadays, the lifespan of a human being is reduced because of heart diseases. So, World Health Organization (WHO) developed targets for prevention of non-communicable diseases (NCDs) in 2013, in which, 25% of relative reduction is from cardiovascular diseases and it is being ensured that at least 50% of patients with cardiovascular diseases have access to relevant drugs and medical counseling by 2025. Around 17.9 million people died just because of cardiovascular diseases in 2016, which is 31% of deaths around the world. A major challenge in heart disease is its detection. It is difficult to predict whether a person has heart disease or not. There are instruments available that can predict heart diseases but either they are expensive or are not efficient to calculate the chance of heart disease in humans.

A check of the World Health Organization (WHO) says that medical professionals can prognosticate just 67 of heart complaints, so there's a vast compass of exploration in this field. In the case of India, access to good croakers and hospitals in pastoral areas is veritably low. 2016 WHO report says that just 58 of the croakers have a medical degree in civic areas and 19 in pastoral areas. Heart conditions are a major challenge in medical

wisdom, Machine Literacy could be a good choice for prognosticating any heart complaint in humans. Heart conditions can be prognosticated using Neural Network, Decision Tree, KNN, etc.

There are many types of Heart disease types based on CP values.

1. Typical Angina is a type of chest pain caused by reduced blood flow to the heart.

2. Atypical Angina pectoris which does not have associated classical symptoms of chest pain.

3. Non Angina or Noncardiac chest pain is often described as feeling like angina, the chest pain caused by heart disease. It feels like a painful squeezing or tightness in your chest, or like pressure or heaviness, particularly behind your sternum. You may feel it on the right side or the left side or in the middle

4. Asymptomatic if a patient tests as a carrier for a disease or infection but experiences no symptoms.

Methodology

Preface World Health Organization has estimated 12 million deaths do worldwide; every time due to Heart conditions. Half the deaths in the United States and other advanced countries are due to cardiovascular conditions. The early prognostic of cardiovascular conditions can prop in making opinions on life changes in high threat cases and turn reduce the complications. This exploration intends to pinpoint the most applicable/ threat factors of heart complaint as well as prognosticate the overall threat using logistic retrogression Data Preparation. Stacking Classifier is a type of retrogression analysis in statistics used for vaticination of outgrowth of a categorical dependent variable from a set of the predictor. In Stacking Classifier, the dependent variable is always double. Logistic retrogression is substantially used for vaticination and also calculating the probability of success.

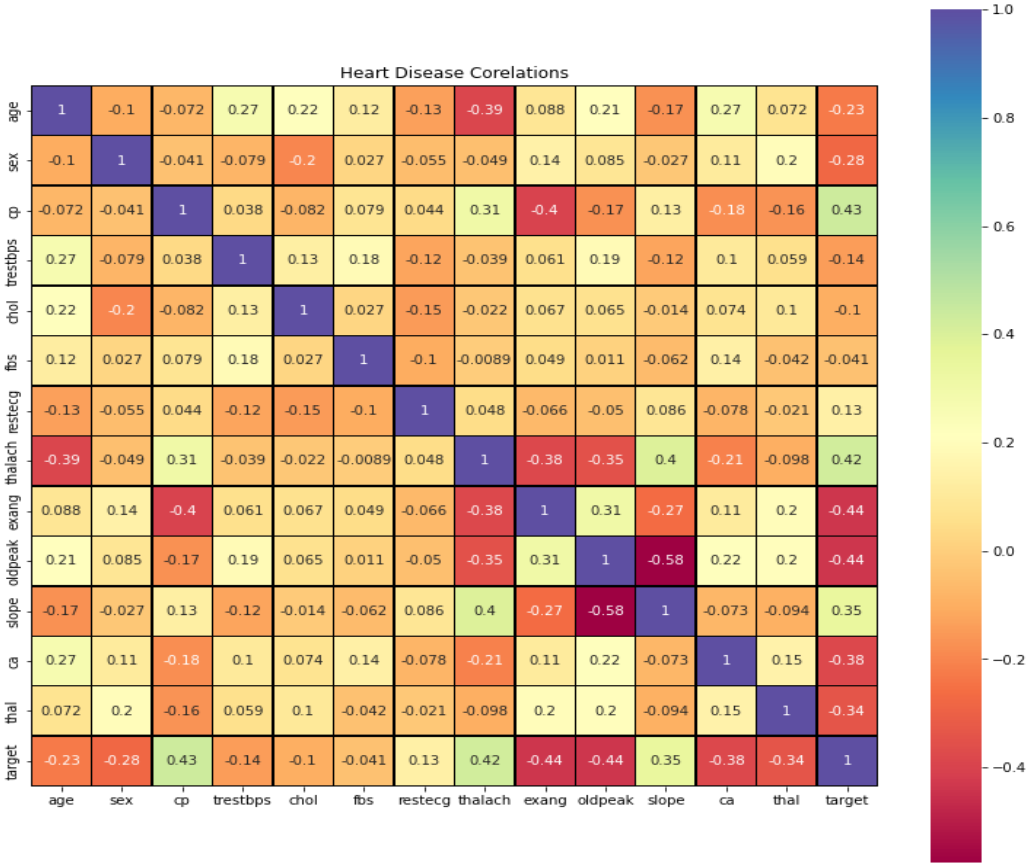
	age	sex	cp	trestbps	chol	fbs	...	exang	oldpeak	slope	ca	thal	target
0	63	1	3	145	233	1	...	0	2.3	0	0	1	1
1	37	1	2	130	250	0	...	0	3.5	0	0	2	1
2	41	0	1	130	204	0	...	0	1.4	2	0	2	1
3	56	1	1	120	236	0	...	0	0.8	2	0	2	1
4	57	0	0	120	354	0	...	1	0.6	2	0	2	1
5	57	1	0	140	192	0	...	0	0.4	1	0	1	1
6	56	0	1	140	294	0	...	0	1.3	1	0	2	1
7	44	1	1	120	263	0	...	0	0.0	2	0	3	1
8	52	1	2	172	199	1	...	0	0.5	2	0	3	1
9	57	1	2	150	168	0	...	0	1.6	2	0	2	1
10	54	1	0	140	239	0	...	0	1.2	2	0	2	1
11	48	0	2	130	275	0	...	0	0.2	2	0	2	1
12	49	1	1	130	266	0	...	0	0.6	2	0	2	1
13	64	1	3	110	211	0	...	1	1.8	1	0	2	1
14	58	0	3	150	283	1	...	0	1.0	2	0	2	1
15	50	0	2	120	219	0	...	0	1.6	1	0	2	1
16	58	0	2	120	340	0	...	0	0.0	2	0	2	1
17	66	0	3	150	226	0	...	0	2.6	0	0	2	1
18	43	1	0	150	247	0	...	0	1.5	2	0	2	1
19	69	0	3	140	239	0	...	0	1.8	2	2	2	1

Figure 1: Dataset Distribution

The dataset as shown in Figure. 1 is from an ongoing cardiovascular study on resides of the city of Framingham, Massachusetts. The bracket thing is to prognosticate whether the case has a 10- a time threat of unborn coronary heart complaint (CHD). The issue has come less clear over recent times, in part, because of misgivings in the attention of heart failure, the lack of methodical recordings of arterial pressure previous to the onset of, and treatment for heart failure, and the absence of methodical visualization. of epicardial coronary highways, that's easily depicted.

Smoking damages the heart and blood vessels veritably snappily, but the damage is repaired snappily for utmost smokers who stop smoking. Indeed, many cigarettes now also damage the heart, so the only proven strategy is to keep your heart safe. From the goods of smoking is to quit. Researchers factors that high blood sugar (glucose) causes stronger contraction of blood vessels and also linked a protein associated with this increased contraction. The results could lead to new treatments to ameliorate issues after a

heart attack or stroke is shown in Figure 5. Age is the most important threat factor in developing cardiovascular or heart conditions, with roughly a tripling of threat with each decade. life. Coronary adipose stripes can begin to form in nonage. It's approximated that 82 percent of people who die of coronary heart complaints are 65 and age. Contemporaneously, the threat of stroke doubles every decade after age 55.



Correlation Analysis: Correlations have three important characteristics. They can tell us about the direction of the relationship, the form (shape) of the relationship, and the degree (strength) of the relationship between twovariables**.**

Figure 2: Correlation Matrix of the Model

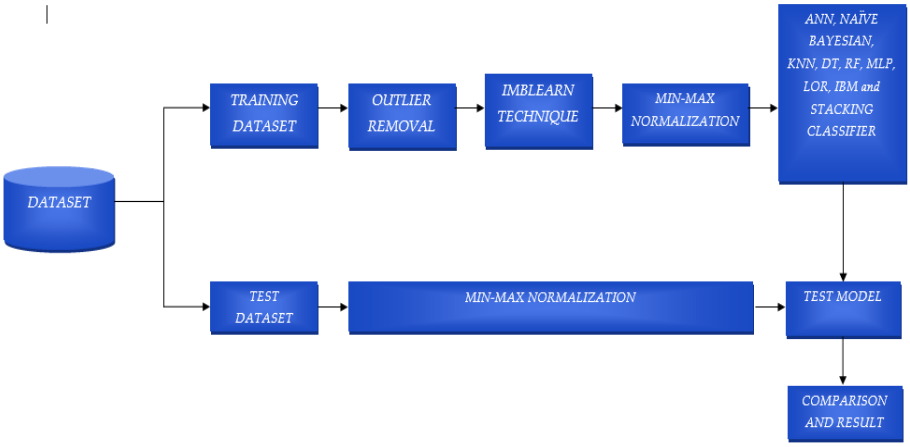


Figure 3: Block Diagram

Figure6 represents the block diagram of proposed system of the heart disease prediction. The block

diagram consists of mainly two steps, the first one is the training dataset and the other one is test data.

1. Naïve Bayes Classification

Naive Bayes classification algorithm tends to be a baseline solution for sentiment analysis tasks. The basic idea of the Naive Bayes technique is to find the probabilities of classes assigned to texts by using the joint probabilities of words and classes. Let's have a brief look at mathematics.

Given the Bayes theorem: $P(A|B) = \frac{P(B|A)}{P(A)}P(B)$

For given elements, A and B and their probability of circumstance P (X) are calculated, where P (A) is the probability of circumstance of element A, P (B) is the probability of circumstance of element A and P (A| B) is the tentative probability of element A given element B occurs, and similar theorem will be used to perform the bracket. So, for independent features, the mentioned theorem would perform a direct addition of the probability of each point passing.

2. Data Set Information

The name of the dataset isheart.csv. There are 303 cases in this dataset, where the cases are either people having heart complaints or they're healthy. Among 303, 165 (54.45) cases are people with a heart complaint and 138 (45.54) are people without heart complaint. The number of attributes is 14. There will be no missing values in the data set nor any null values.

Features include age, coitus, casket- pain type, rest BP, cholesterol, blood sugar position, ECG result, maximum heart rate achieved, exercise-convinced angina, ST depression, the pitch of peak exercise ST member, number of major vessels, and disfigurement in heart as of 3-normal, 6- fixed disfigurement and 7-reversible disfigurement.

3. KNN Methods

K-Nearest Neighbor (K-NN) In the K-NN algorithm a data point is taken whose classification is not available, then the number of neighbors, k is defined. After that k neighbors are selected according to the lowest Euclidiandistance between the selected data points and their neighbors. The selected data point is then classified into a category, which is the same as the category which has the majority of neighbors among the K neighbors.

4. Random Forest Methods

Random Forest works by building multiple decision trees of the training data. each of the trees predicts a class as an affair and the class, which is the affair of the topmost number of decision trees is taken as the final result, in the case of the bracket. In this algorithm, we need to define the number of trees we want to produce. Random Forest is a bootstrap aggregating or bagging fashion. This fashion is used to drop the friction in the results.

All classification and regression are achieved utilizing Random Forest algorithms. The information is categorized into a tree, and predictions are based on that tree. With a significant number of record values missing, the Random Forest algorithm still can produce the same results when deployed to huge datasets. The decision tree's generated samples can be preserved and used for different data sets. In a random forest, there are two stages: first, generate a random forest, and then, using a classifier produced in the first stage, make a prediction.

5. Decision Tree

The Decision Tree algorithm's central node symbolizes the dataset traits, while the outlying branches will accomplish a specific objective. Decision trees are used because they are quick, reliable, simple to understand, and require very little data preparation. The decision tree's root determines the class label prediction. The core attribute's value, as well as the record's attribute, are evaluated.

6. Stacking Classifier

Stacking is a way of ensembling classification or regression models it consists of two-layer estimators. The first layer consists of all the baseline models that are used to predict the outputs on the test datasets. The second layer consists of Meta-Classifier or Regressor which takes all the predictions of baseline models as an input and generate new predictions.

Specifically, we will evaluate the following five algorithms:

- Multi-Layer Perceptron
- Adaboost Classifier
- Decision Tree.
- Extra Tree Classifier

Results

Our main goal going into this research was to predict Heart disease using various machine learning techniques. We predicted using K-Nearest Neighbor (K-NN) , ANN, DT, RF, IBM, MLP,LOR and Stacking Classifier. Stacking classifier gives 99.96 % of accuracy with better results. With Each algorithm, we have observed Accuracy, Precision, Sensitivity and Specificity as follows:

```
STACKING CLASSIFIER Accuracy is :99.00%
```

```
from sklearn.metrics import classification_report
STK_Pred=STK.predict(x_test)
STKreport = classification_report(y_test, STK_Pred)
print(STKreport)
```

```
              precision    recall  f1-score   support

           0       0.98      1.00      0.99       117
           1       1.00      1.00      1.00        33
           2       0.98      0.97      0.98        66
           3       1.00      0.94      0.97        17

    accuracy                           0.99       233
   macro avg       0.99      0.98      0.98       233
weighted avg       0.99      0.99      0.99       233
```

Figure 4: With SMOTE and CP Based STACKING CLASSIFIER

```
confusion matrix for STACKING CLASSIFIER :
<sklearn.metrics._plot.confusion_matrix.ConfusionMatrixD
```

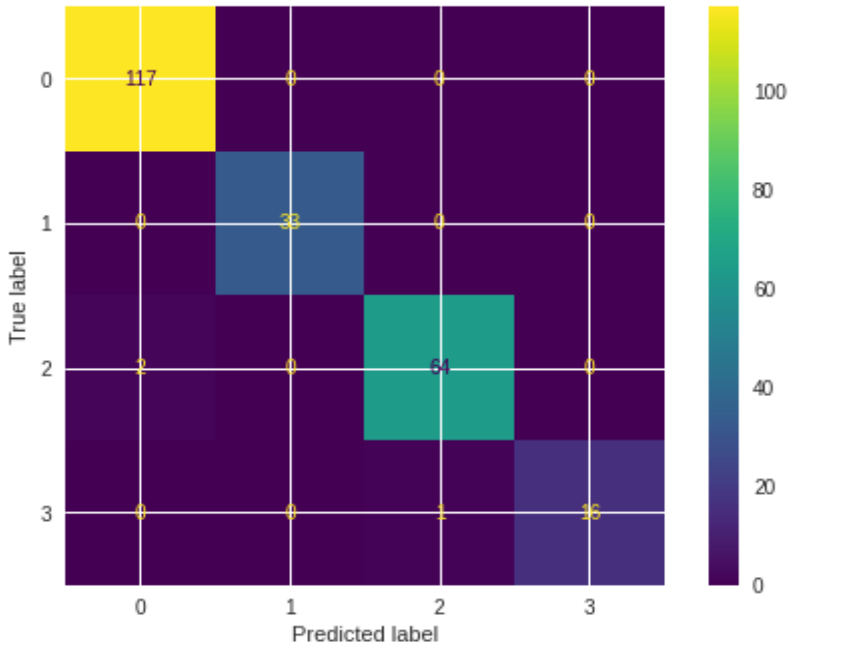


Figure 5: With SMOTE and CP Based Confusion Matrix

II. Comparison Chart

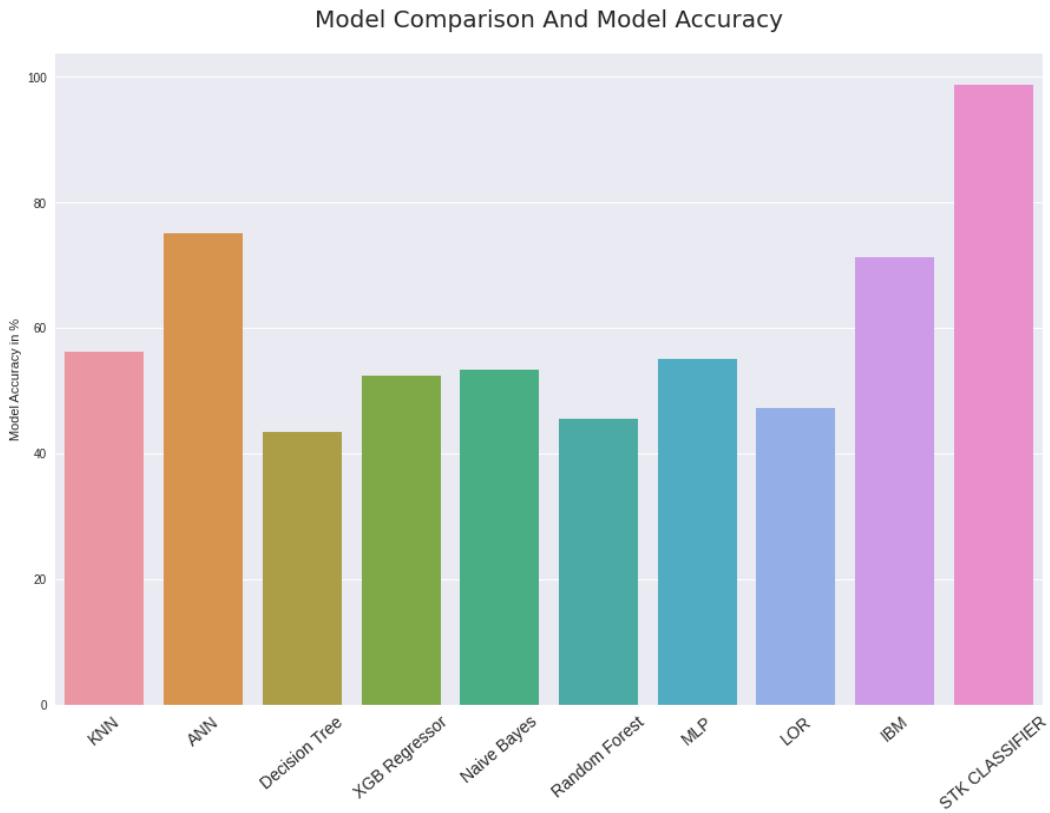


Figure 6: With SMOTE and CP Based Comparison Chart

III. Conclusion

The number of Heart conditions can exceed the current script to reach the maximum point. Heart complaint Is complicated and each and every time lots of people are dying from this complaint. It's delicate to Manually determine the odds of getting a heart complaint grounded on threat factors preliminarily shown. By Using this system one of the major downsides of this work is that its main focus is aimed only at the operation of classifying ways and algorithms for heart complaint vaticination, by studying Colorful data cleaning and mining ways that prepare and make a dataset applicable for data Mining so that we can use this Machine Learning in that logistic retrogression algorithms by prognosticating If the case has heart complaint or not. The any-medical hand can use this software and prognosticate heart complaints and reduce the time complexity of the croakers. It's still an open sphere staying to get Enforced in heart complaint vaticination and increase the delicacy. After applying colorful algorithms, it can be said that machine literacy is proving to be extremely precious in prognosticating heart complaints which is one of the most prominent problems of society in the moment's world. As further and further work is being done in the field of machine literacy, soon there may be new styles to make machine learning further helpful in the field of healthcare. The algorithms used in this trial have performed well using the available attributes. The conclusion can be eventually drawn that machine literacy can reduce the damage done to a person physically and mentally, by prognosticating heart complaints.so in this research using Stacking Classifier model with SMOTE and CP value that resulted in highest accuracy with an accuracy of 99% predict the accuracy and gives 91% of accuracy using Stacking Classifier model without SMOTE and CP value to identify the Heart Disease risk based on CP value.

IV. References

[1] Avinash Golande, Pavan Kumar T. Heart complaint vaticination using effective machine Learning ways.

[2] The Lancet Global Health. The changing patterns of cardiovascular conditions and their threat factors in the countries of India The global burden of complaint study 1990-2016.

[3] Himanshu Sharma, M A Rizvi. Vaticination of heart complaints using machine literacy algorithms A check. World health ranking.

[4] Sana Bharti, Analytical study of heart complaint vaticination compared with different algorithms; International conference on computing, communication, and robotization (ICCA2022).

[5] Monika Gandhi, Vaticination in heart complaint using ways of data mining, international conference on futuristic trend in computational analysis and knowledge operation (ABLAZE-2022)

[6] Sarath Babu, 2017. heart complaint opinion using data mining fashion, international national conference on electronics, communication and aerospace technology (ICECA2022)

[7] A H Chen, 2011. HDPS heart complaint vaticination system;

[8] Grover SA, Lowensteyn I, Esrey KL, et al. Do croakers directly assess the coronary threat in their cases. Primary results of the Coronary Health Assessment Study.

[9] Friedmann PD, Brett AS, Mayo-Smith MF. Differences in generalists and cardiologists' comprehensions of cardiovascular threat and the issues of preventative remedy in the cardiovascular complaint.

[10] Meland E, Laerum E, Lehmann EH. Assessment of coronary heart complaint threat, II a clinical multi- center study of general interpreters threat assessment.

[11] Mohan.S, Thirumalai.C, & Srivastava.G, (2021). Effective heart complaint vaticination using modern machine-learning ways.