

Prediction of Cardiac Disease Based on Patient's Symptoms

I Juvanna^a, R VinithVaran^b, R Ragul^c, G B Bharath Raj^d

^{a,b,c,d} Department of Information Technology, Hindustan Institute of technology and science, Chennai

^a ijuvanna@hindustanuniv.ac.in, ^b vinithravi005@gmail.com, ^d saibharathraj23@gmail.com, ^c ragulflamer04@gmail.com

Article History Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

Abstract: The detection of comparative sign of likely side effects and symptoms which may additionally rapid identification of cardiovascular diseases utilizing data gathered from past patients just as input data taken from client at that specific time. Recent condition of healthcare system data scrutiny for observation are not, at this point fundamentally a period building arrangement of daily counts. All things considered, a profusion of proposed longitudinal just as temporal demographic, symptom data are accessible at data introduced during the time of execution. Our proposed method contains all data that is being exploited as classification approach that analyses recent health-care data against data from that specific base distributed and hence classifies subclasses the given data. Likewise, test data utilized is tried against different kinds of classification, other propose test scores have been of prediction.

Keywords: Classification accuracy comparison, Disease diagnosis, Prediction, Decision tree

1. Introduction

Major deadly assaults on the planet are heart failure which leads to uniqueness of death. Coronary failures are achieved by a sudden occasion of heart apoplexy, usually which result about the end of a heart muscle and portion of the time can be lethal [1]. A heart attack happens if progression of oxygen-rich blood to section of heart muscle all of a sudden becomes obstructed and heart cannot get oxygen. Exactly when plaque creates in corridors, condition is called atherosclerosis. The master creates of plaques present in stockpile courses occur over numerous years. In the end, a region of plaque can burst (tear open) inside a vein. This can cause a blood coagulation that can outlined on a surface level of the plaque [2]. The progression of blood through the heart corridor gets completely hindered once square ends up being gigantic. If any blockage present isn't dealt with quickly, the fragment of heart muscle that can be dealt with by corridor can be provoke passing of that particular course. sound heart tissue is displaced with scar tissue. This heart damage may not act naturally apparent, which further caused tough and outrageous issues [3]. A larger piece of heart attack sharpens as needs be degree to heart attacks. The heart sicknesses are coronary supply route gets totally obstructed once the clot turns out to be huge. In the event that any blockage present isn't a condition where a wax like substance that can be named as plaque creates inside the heart arteries. Simply early expectation could help with bettering investigate a coronary failure toward early phase to save a person life. A less begun essential justification coronary episode is that of outrageous fit or rather fixing of a heart vein [4]. Atherosclerosis doesn't show sway on fits present in the heart supply routes. Coronary episode that can connected with or can incite outrageous issues that can decrease the strength of a person, for instance, cardiovascular failure and besides can provoke perilous arrhythmias. Coronary failure is a condition wherein heart can't siphon sufficient blood to address the body issues. Flighty heartbeats are called arrhythmias [5]. Ventricular fibrillation present, related with perilous arrhythmia that prompted demise if not treated advantage away. There are couple of components that can impact a person tendency for cardiac illnesses. Guidance is a huge pointer of monetary status that is connected with occupation and besides among various parts affecting a individual lifestyle [6]. Different packs in caused nations to have shown that Cardiac sicknesses rate fluctuates between individuals with various degrees of direction. Data burrowing procedure is inserted for the use of instruments that can oversee complex data assessment to find successfully dull, liberal models and affiliations present in this colossal dataset. These mechanical assemblies that can intertwined with the given models that consolidates quantifiable figures, mathematical computations, machine proposed learning techniques will impel early malignant growth of illness [7]. acknowledging which fuses depiction of these specific models, this learning plan is introduced close by a specific strategy of mentioned models especially which can be seen as conventional what's more will be set up to learn in a method for broadening this specific depiction of covered models. In related learning, any connection that among the features given is looked for, not only for the ones that expect in a specific given class regard. Especially parties that fuse bundling are fused for models which have a spot together are looked for. In the presumption for numeric characteristics, the improved result that can be assumption isn't only a discrete class yet this likewise goes presumably as a mathematical aggregate. In this assessment, which consolidates the solicitation for that specific data which can be picked for that infrequently utilized models that are connected with the dataset by decision tree figuring [8]. Decision tree that acts like flowchart outline like plan,

where the present every inside center that can be settled and inferred a test which fuses a trademark, each branch presents watches out for a result of that specific test, each leaf center point that can incorporate moreover, hold the class mark. The specific center which is available at pinnacle is root center point. The quality assessment of that specific data endeavored what's more, checked against a decision tree. A way which can be followed from that specific current root center to the leaf center, can hold and avow the normal class for that specific data. Decision tree which can without a doubt changed over into the given game plan of guidelines being referred to [9]. Choice tree which is used to makes the unremitting models in dataset can take that information and other thing sets that happen a significant part of the time in the informational index which are known as continuous examples. Information grouping is the communication that incorporates relationship of information into particular arrangements for its best and gainful use. A particularly masterminded grouping of information arrangement framework simplifies essential information to find, recuperate. This can be that of particular importance of hazard the board, legitimate revelation, and consistence. At the point when an arrangement to bunch the information will be made, the security norms that demonstrate fitting managing practices for each class classification, principles of limit that characterize lifecycle necessities of the information ought to be tended to [10].

2. Related Work

The scientists that utilizes design acknowledgment of information mining strategies helps in anticipating models dependent on the cardiovascular analyze area. The trials that were completed utilizing these order-based calculations, for example, Naïve Bayes, Decision tree, K-NN, Neural network these outcomes has demonstrated to that of Naïve Bayes strategy that have performed better than compared to the other when used through the methods. The scientists utilize K methods grouping calculation on that specific coronary illness where the stockroom which identify with extricate information pertinence to the coronary illness, applies to type MAFIA (Maximal Frequent Item set Algorithm) calculations to compute weightage of continuous examples are presumably important to respiratory failure forecasts [11].

Numerous scientists had proposed a layered engineering of neuro-fluffy methodology which can be assist with foreseeing the events or identification regarding coronary illness recreated utilizing MATLAB instrument. The execution of this neuro-fluffy incorporated methodology that assists us with creating a consequence of a mistake rate at an exceptionally low and high work productivity in performing examination for the coronary illness events. The specialists additionally proposed another path approach that affiliation rule mining dependent on given grouping number, bunching value-based informational index for a recommended indication for the event of coronary illness expectation. The execution of this propose approach was carried out in JAVA programming which decreased the fundamental memory prerequisite by thinking about a little bunch at some random time in order to be viewed as truly adaptable and productive enough.

The specialists have utilized the information mining calculations like hereditary calculation, guileless Bayes, affiliation characterization, choice trees, and neural organizations for forecast and dissecting the consequences of coronary illness from given dataset. A trial performed by the scientist proclaimed that on given dataset when the model is created utilizing cross breed wise calculation and neural organizations the results show that mixture strategy improved exactness of the given expectations.

The exploration paper depicted above recounts model that are being utilized i.e., weighted acquainted classifier (WAC), Naïve Bayes to foresee likelihood of the quantity of patient getting coronary failures [10][16]. The specialists shows that advancement for a wise framework which was online utilizing guileless Bayes calculation to furnish the appropriate response with an unpredictable question for diagnosing the coronary illness and helps clinical professionals with given clinical choices. The scientist was utilizing the given affiliation decides that assists us with addressing a specific strategy in information mining that assists with improving the illness forecast which can be furnished with extraordinary potential and sufficient exactness. A calculation with which this inquiry requirements can work was additionally acquainted with help coronary corridor gets totally hindered once the coagulation turns out to be huge. In the event that any blockage present is not decreasing quantity of affiliation governs and anticipate the approved worth utilizing train& test approach. A three famous information mining calculations which be assists us with supporting an vector machines, counterfeit Neural organizations, followed by the choice tree was related by the specialists to build up potential expectation model utilizing an 502 cases. SVM got outstanding amongst other forecast models followed by counterfeit neural organizations.

The specialist's expectation which utilizes choice trees, innocent Bayes, and neural organization to foresee the potential indications for coronary illness with 15 mainstream ascribes as hazard factors recorded in the clinical writing.

Spreading the models acquired with two sub-divisions of information mining calculations had been named developmental, false names specifically GA-KM and MPSO-KM group of cardiovascular infection informational

collection& anticipate model precision. This is a cross breed strategy that joins bunching implies method with energy type molecule swarm advancement (MPSO) for better outcomes. A given examination shown is being made in field of exploration directed utilizing C5, Naïve Bayes, K-implies, Ga-KM , MPSO-KM for arrangement of conveying exactness from grave methods. This trial result validated above outcomes with improve exactness and upgrades it when utilizing GA-KM, MPSO-KM [12].

Analysts made class affiliation rule which includes the subset determination to foresee a model of coronary illness. The forecast for Association rule assists with deciding the relations anticipated among the ascribes qualities and characterization predictions of class in the patient dataset. Highlight choice estimates like the hereditary pursuit technique assists with deciding attributes which contribute towards expectation for heart sicknesses. The analysts executed a crossover framework that utilizes worldwide enhancement advantage of hereditary calculation that specifically helps us for instatement of these neural organization loads. The expectation of the coronary illness depends on hazard factors like family, diabetes, age, history, hypertension, elevated cholesterol, smoking, liquor admission , heftiness.

Machine Learning

AI is an upward innovation which permits PCs to gain naturally from past information. AI customs different calculations for building numerical models and making forecasts utilizing old information or data. As of now, it is being utilized for different errands, for example, picture acknowledgment, discourse acknowledgment, email sifting, Facebook auto-labeling, recommender framework, and some more.

AI is asserted as a development of man-made reasoning that is to a great extent worried about the development of calculations which grant a PC to gain from the information and past proficiencies all alone. The term AI are first introduced by the Arthur Samuel in 1959. We will characterize it in a summed up manner as: "AI empowers a machine to consequently gain from information, improve show from encounters, and anticipate things without being unequivocally modified".

A Machine Learning framework gains from authentic information, fabricates the forecast models, and at whatever point it gets new information, predicts the yield for it. The precision of anticipated yield relies on the measure of information, as the huge measure of information assists with building a superior model which predicts the yield all the more precisely. Assume we have a diverse issue, where we need to accomplish a few forecasts, so as opposed to composing a code for it, we simply need to provender information to generic algorithms, with the assistance for these calculations, machine assembles rationale according to information, foresee the yield. AI has changed our perspective about this issue.

Classification of Machine Learning

A.Supervised Learning

Supervised learning is regularly characterized as learning with correct regulator else you can say that learning inside present of educator. The algorithm learns on a labelled dataset with an answerkey and does the training and evaluation. Administered learning is anticipated on "train me" idea. Supervised learning has next measures:

- Classification
- Random forest
- Decision tree
- Regression

There are following machine AI algorithms :

- Linear Regression
- Logistical Regression
- Support Vector Machines (SVM)
- Neural Networks
- Random Forest
- Gradient Boosted Trees
- Decision Trees
- Naive Bayes

B. Unsupervised learning

Unsupervised learning is regularly characterized as the learning without an authority which in Unsupervised realizing there are no instructor are directing. Unsupervised realizing when a dataset is given it precisely work on dataset , discover example, connection among them and understanding to the framed connections, when new information is given it group them and store in one of them connection. Solo learning depends on "independent " idea. For instance, assume there are blend natural products mango, banana and apple , when Unsupervised learning is applied it order them in three distinct bunches on the premise if their connection with one another and when another information is given it consequently send it to one of the groups. Supervisor learning say there are mango, banana, apple however Unsupervised learning said it as there are three distinct clusters. Unsupervised learning had following interaction:

- Dimensionality
- Clustering

There are following unaided AI calculations:

- t-SNE
- k-implies bunching
- PCA

C. Reinforcement

Reinforced learning is specialised ability to help out climate, discover an result. It depends on "hit and preliminary" idea. In built up learning every specialist is offered with positive and negative focuses and based on certain focuses reinforced learning give dataset yield that is based on sure distinctions it prepared and based on this preparation play out the testing on datasets.

3. Existing System

This framework, the info subtleties are obtained from the patient. At that point from the client inputs, utilizing ML methods coronary illness is analyzed [13]. Presently, acquired outcomes are connected with after effects of existing models inside a similar area, discovered to be overhauled. The information of coronary illness patients gathered from the UCI research facility is utilized to learn designs with NN, DT, Support Vector machines SVM, Naive Bayes [14]. The proposed hybrid technique returns consequences of 86% for F-measure, testing with other existing strategies.

4. Proposed System

In the wake of assessing the outcome from the overarching strategies, we've utilized python & pandas tasks to perform heart condition order for information acquired from the UCI repository. It gives a simple to-utilize visual delineation of the dataset, working climate and building the prescient examination. ML measure begins from a pre-handling information stage followed by highlight determination dependent on information cleaning, and train the framework utilizing CNN, grouping of displaying execution assessment. Random forest procedure is utilized to improve precision for the outcome.

5. Methodology

- Dataset Collection
- Data Pre-Processing
- Split the Data Train and Test
- Model Fitting
- Testing
- Evaluate

6. Analysis

6.1 Data Pre-Processing

Cardiovascular sickness information is pre-prepared by utilizing different assortment of accounts. A dataset holds a sum of 895 patient records in Figure 1. where 6 records are for certain missing qualities. Those 6 records have been separated from dataset , excess 889 patient records are utilized in pre-preparing in Figure 2.

```
print(train_data.head(10))
```

PatientID	Gender	Chain_smoker	Tobacco_products	HighBP	Obesity	Diabetes	Metabolic_syndrome	Stimulant_drugs	Family_history	History_of_preeclampsia	CABG_history	Respiratory_illness	UnderRisk
0	1	1	1	0	1	0	0	0	1	0	0	0	no
1	1	0	1	0	1	0	0	0	1	0	0	0	no
2	1	0	1	0	1	0	0	0	1	0	0	0	no
3	1	0	1	0	1	0	0	0	1	0	0	0	no
4	1	0	0	0	0	0	1	1	0	0	0	0	no
...
95	1	0	1	0	1	0	0	0	1	0	0	0	no
96	1	0	0	0	1	0	0	0	1	0	0	0	no
97	1	0	1	0	1	0	0	0	1	0	0	0	no
98	1	1	1	1	1	0	0	0	1	0	0	0	no
99	2	0	1	1	1	0	0	0	1	0	0	1	yes

100 rows × 14 columns

Figure1. Dataset

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 889 entries, 0 to 888
Data columns (total 14 columns):
#   Column                                Non-Null Count  Dtype
---  ---                                -
0   PatientID                             889 non-null    int64
1   Gender                                 889 non-null    int64
2   Chain_smoker                          889 non-null    int64
3   Tobacco_products                      889 non-null    int64
4   HighBP                                889 non-null    int64
5   Obesity                                889 non-null    int64
6   Diabetes                               889 non-null    int64
7   Metabolic_syndrome                   889 non-null    int64
8   Stimulant_drugs                       889 non-null    int64
9   Family_history                        889 non-null    int64
10  History_of_preeclampsia               889 non-null    int64
11  CABG_history                          889 non-null    int64
12  Respiratory_illness                   889 non-null    int64
13  UnderRisk                             889 non-null    object
dtypes: int64(13), object(1)
memory usage: 97.4+ KB
```

Figure 2. pre-processing

6.2. Train and Test the Model

This stage is to make assessment the models upheld the info record. For our motivation of study, we are having the chance to execute to mentor the model utilizing CNN algorithm in thick layer. Thick layer is that normal profoundly associated neural organization layer. It is commonest and routinely utilized layer. Thick layer does the underneath procedure on the information & return the yield

```
test_data.describe()
```

	PatientID	Gender	Chain_smoker	Tobacco_products	HighBP	Obesity	Diabetes	Metabolic_syndrome	Stimulant_drugs	Family_history	History_of_preeclampsia	CABG_history	Respiratory_illness
min	362.000000	362.000000	362.000000	362.000000	362.000000	362.000000	362.000000	362.000000	362.000000	362.000000	362.000000	362.000000	362.000000
max	291.500000	1.250000	0.125000	0.750000	0.650000	0.850000	0.850000	0.850000	0.850000	0.850000	0.850000	0.850000	0.850000
min	116.416146	0.519127	0.325070	0.405547	0.230220	0.309950	0.230186	0.309950	0.260271	0.672262	0.143379	0.171462	0.000000
max	161.250000	1.000000	0.000000	1.000000	0.000000	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
min	291.500000	1.000000	0.000000	1.000000	0.000000	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
max	362.000000	2.000000	0.000000	1.000000	0.000000	1.000000	0.000000	0.000000	1.000000	0.000000	0.000000	0.000000	0.000000
min	442.000000	2.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000

6.3. Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA), essentially a sort of story for analysts. It licenses us to find examples and experiences, regularly with visual strategies, inside information. EDA is frequently the initial step of the information demonstrating measure.

Information imagine for both Training and testing Datasets for all highlights utilizing EDA ideas.

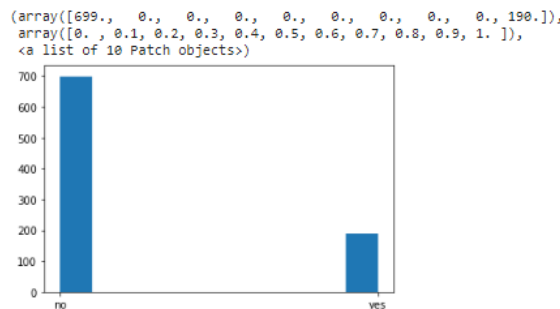


Figure 3. Risk analysis

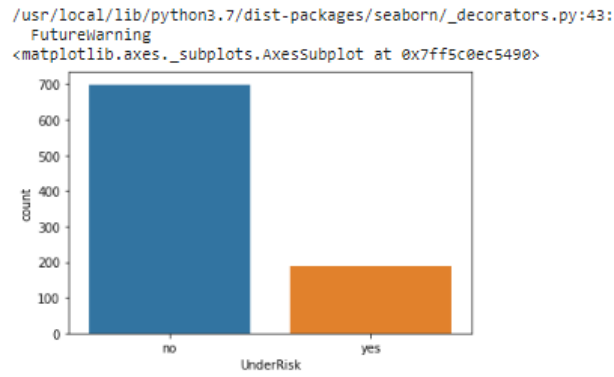


Figure4. Under Risk Count

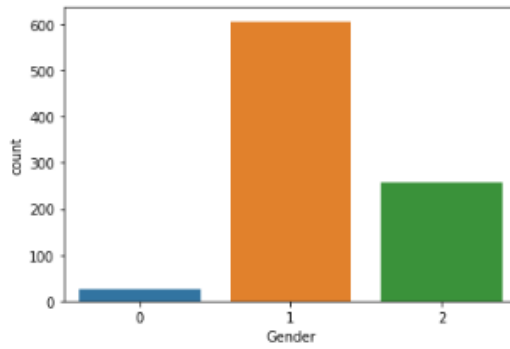


Figure 5. Gender count analysis

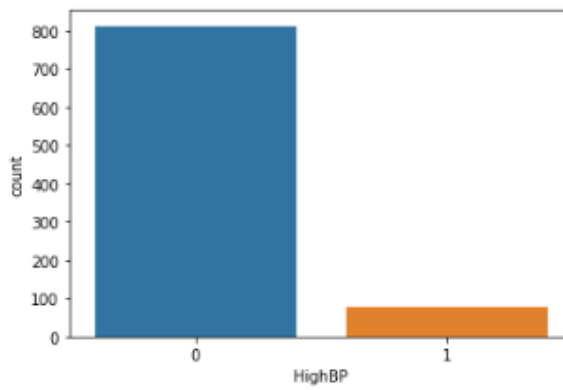


Figure 6. High BP Patient count Analysis

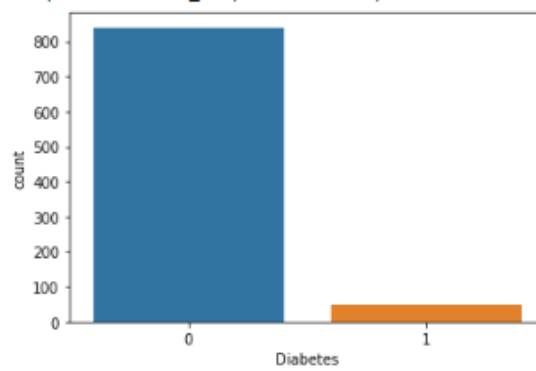


Figure 7. Diabetes patients count analysis

6.4. CONVOLUTIONAL NEURAL NETWORK (CNN):

Input Layers:

The quantity of neurons in this layer is equivalent to sum of number of features(input) in this information .

Hidden Layer:

The contribution from input layer is then fed into the secret layer, which can be many , relying on our model and information size . Each secret layer can be various quantities of neurons which are by and large more noteworthy than quantity of highlights. Output from each layer is processed by the framework duplication of yield of the past layer with learnable loads of that layer and afterward by expansion of learnable inclinations followed by initiation work which makes the organization nonlinear in Figure 3.

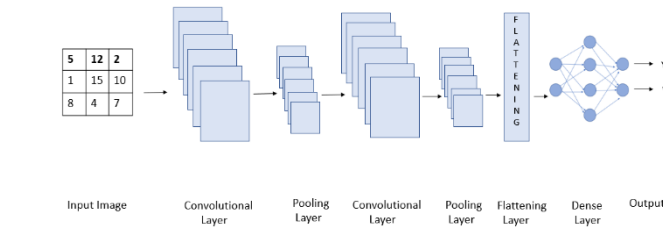


Figure8. Convolutional Neural Network architecture

Output Layer:

The output from the secret layer is then fed into a strategic capacity comparative sigmoid or SoftMax which changes over the yield of each class into probability score of each class in Figure 4.

```

Epoch 1/50
6/6 [=====] - 1s 77ms/step - loss: 0.7172 - accuracy: 0.5159 - val_loss: 16.0429 - val_accuracy: 0.8202
Epoch 2/50
6/6 [=====] - 0s 8ms/step - loss: 0.7164 - accuracy: 0.5514 - val_loss: 9.1564 - val_accuracy: 0.8202
Epoch 3/50
6/6 [=====] - 0s 9ms/step - loss: 0.6971 - accuracy: 0.5941 - val_loss: 5.5919 - val_accuracy: 0.8202
Epoch 4/50
6/6 [=====] - 0s 11ms/step - loss: 0.6812 - accuracy: 0.6275 - val_loss: 4.3899 - val_accuracy: 0.8202
Epoch 5/50
6/6 [=====] - 0s 9ms/step - loss: 0.6781 - accuracy: 0.6261 - val_loss: 4.1118 - val_accuracy: 0.8202
Epoch 6/50
6/6 [=====] - 0s 8ms/step - loss: 0.6400 - accuracy: 0.7114 - val_loss: 5.1567 - val_accuracy: 0.8202
Epoch 7/50
6/6 [=====] - 0s 11ms/step - loss: 0.6527 - accuracy: 0.6988 - val_loss: 9.1368 - val_accuracy: 0.8202
Epoch 8/50
6/6 [=====] - 0s 8ms/step - loss: 0.6564 - accuracy: 0.6829 - val_loss: 0.6239 - val_accuracy: 0.8202
Epoch 9/50
6/6 [=====] - 0s 9ms/step - loss: 0.6369 - accuracy: 0.6962 - val_loss: 8.1014 - val_accuracy: 0.8202
Epoch 10/50
6/6 [=====] - 0s 9ms/step - loss: 0.6385 - accuracy: 0.7261 - val_loss: 8.6889 - val_accuracy: 0.8202
Epoch 11/50
6/6 [=====] - 0s 11ms/step - loss: 0.6145 - accuracy: 0.7357 - val_loss: 7.4179 - val_accuracy: 0.8202
Epoch 12/50
6/6 [=====] - 0s 9ms/step - loss: 0.6012 - accuracy: 0.7365 - val_loss: 6.8202 - val_accuracy: 0.8202
Epoch 13/50
6/6 [=====] - 0s 9ms/step - loss: 0.6097 - accuracy: 0.7264 - val_loss: 5.9995 - val_accuracy: 0.8202
Epoch 14/50
6/6 [=====] - 0s 9ms/step - loss: 0.6018 - accuracy: 0.7391 - val_loss: 5.5434 - val_accuracy: 0.8202
Epoch 15/50
6/6 [=====] - 0s 9ms/step - loss: 0.5964 - accuracy: 0.7435 - val_loss: 5.2013 - val_accuracy: 0.1798
Epoch 16/50
6/6 [=====] - 0s 9ms/step - loss: 0.5868 - accuracy: 0.7533 - val_loss: 3.1133 - val_accuracy: 0.1798
Epoch 17/50

```

Figure9. Training dataset

6.5. XG Boost Algorithm

XGBoost is execution of Gradient Boosted decision trees. This library was written in C++. it's a sort of Software library that was planned fundamentally to improve speed and model execution. It's as of late been ruling in applied AI. XGBoost models significantly overwhelm in numerous Kaggle Competitions [15].

In this calculation, decision trees are made in successive structure. Loads assume a critical part in XGBoost. Loads are relegated to all or any the autonomous factors which are then taken care of into the decision tree which predicts results. Weight of factors anticipated wrong by the tree is expanded and these the factors are then taken care of to the subsequent choice tree in Figure 5. These individual classifiers/indicators then troupe to offer a powerful and more exact model. It can chip away at regression, order, positioning, and client characterized forecast issues.

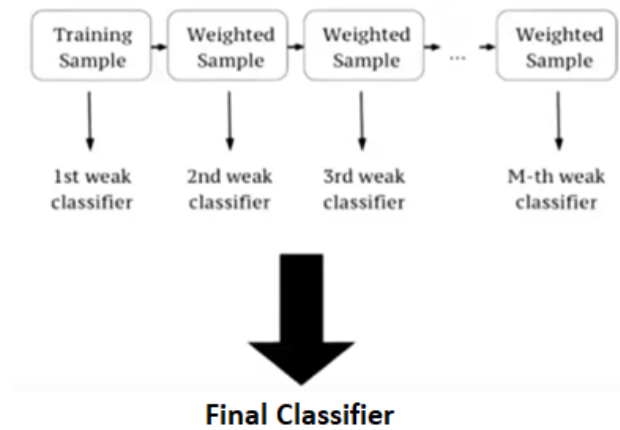


Figure10. XGBoost classifier architecture

7. Performance Measures

A few standard introduction measurements like exactness. Exactness, error in characterization have been considered for calculation of execution viability of this model in Figure6.

Support vector machine (SVM) algorithm was used for same model. The accuracy rate was below 80 percent, while our proposed model has an accuracy more than 90 percent, also SVM takes lot more time to process huge data when compared to CNN.

We compared CNN with SVM algorithm, CNN is more suitable and gives more accurate and, especially for the model that we have proposed. CNN can process large amount of data sets with its hidden layers.

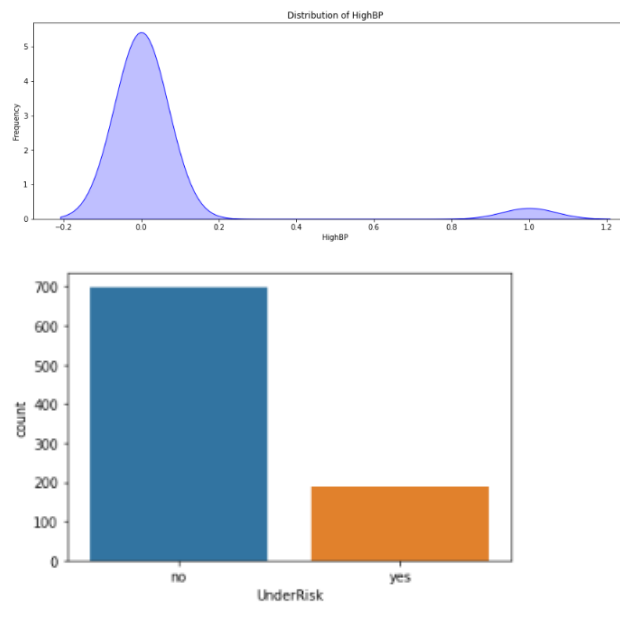


Figure 11. GRAPH ACCURACY (CNN)

8. Conclusion

In this paper, we anticipated a strategy for Cardiac illness expectation utilizing AI and Deep Learning strategies, these outcomes introduced an extreme precision standard for creating a superior assessment result. By presenting new proposed XG Boost order, we track down the hazardous of expectation rate without gear and propose a way to deal with assessment the pulse and condition. we discover the data from the above input through ML Techniques. Initially, we acquainted Deep Learning Perception with train the Model utilizing CNN calculation on datasets

References

- [1]Baitharu, Tapas Ranjan, Subhendu Kumar pani. "Analysis of data mining techniquesfor healthcare decision support system using liver disorder dataset." *ProcediaComputer science* 85 (2016): 862-870.
- [2]Gavhane, Aditi, Gouthami Kokkula, Isha pandya, Kailas Devadkar. "Prediction of heart disease using machine learning." *Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pp. 1275-1278. IEEE, 2018.
- [3]Singh, Smriti Mukesh, and Dinesh B. Hanchate. "Improving disease prediction by machine learning." *Int J Res Eng Technol* 5, no. 6 (2018): 1542-1548.
- [4]Taneja Abhishek. "Heart disease prediction system using data mining techniques." *Oriental journal of computer science and technology* 6,no.4(2013):457-466.
- [5]Dangare, Chaitrali S., Sulabha S. Apte. "Improved study of heart disease prediction system using data mining classification techniques." *International Journal of Computer Applications* 47, no. 10 (2012):44-48
- [6]Thomas, J, and R. Theresa Princy. "Human heart disease prediction system using data mining techniques." *International conference on circuit, power and computing technologies (ICCPCT)*, pp. 1-5. IEEE, 2016.
- [7]Kaur, Beant, and Williamjeet Singh. "Review on heart disease prediction system using data mining techniques." *International journal on recent and innovation trends in computing and communication* 2, no. 10 (2014): 3003-3008.
- [8]Meyer, Alexander, Dina Zverinski, Boris Pfahringer, JörgKempfert, Titus Kuehne, Simon H. Sündermann, Christof Stamm, Thomas Hofmann, Volkmar Falk, and Carsten Eickhoff. "Machine learning for real-time prediction of complications in critical care: a retrospective study." *The Lancet Respiratory Medicine* 6, no. 12 (2018): 905-914.
- [9]Rajkomar, Alvin, Michaela Hardt, Michael D. Howell, Greg Corrado, and Marshall H. Chin. "Ensuring fairness in machine learning to advance health equity." *Annals of internal medicine* 169, no. 12 (2018): 866-872.
- [10]Rajamhoana, S. P., C. Akalya Devi, K. Umamaheswari, R. Kiruba, K. Karunya, and R. Deepika. "Analysis of neural networks based heart disease prediction system." *11th International Conference on Human System Interaction (HSI)*, pp. 233-239. IEEE, 2018.
- [11]Ramalingam, V. V., AyantanDandapath, and M. Karthik Raja. "Heart disease prediction using machine learning techniques: a survey." *International Journal of Engineering & Technology* 7, no. 2.8 (2018): 684-687.
- [12]Kohli, Pahulpreet Singh, Shriya Arora. "Application of machine learning in disease prediction." *4th International conference on computing communication and automation (ICCCA)*, pp. 1-4, IEEE, 2018.
- [13]Marimuthu, M., M. Abinaya, K. S. Hariesh, K. Madhankumar, and V. Pavithra. "A review on heart disease prediction using machine learning and data analytics approach." *International Journal of Computer Applications* 181, no. 18 (2018): 20-25.
- [14]Beyene, Chala, Pooja Kamat. "Survey on prediction and analysis the occurrence of heart disease using data mining techniques." *International Journal of pure and Applied Mathematics* 118, no.8(2018): 165-174.
- [15]Khourdifi, Youness, Mohamed Bahaj. "Heart disease prediction and classification using machine learning algorithms optimized by particle swarm optimization and ant colony optimization." *International Journal of Intelligent Engineering & Systems* 12, no. 1 (2019): 242-252