IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 3, March 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 3, March 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1103077|

Heart Disease Prediction Using Machine Learning

M.Nirmala[1], A.Chandra Sekhar Reddy[2], Ch.Maheswari[3], V.Keerthi Reddy[4], Ch.Shamitha[5]

Assistant Professor, Department of Information Technology, Kallam Haranadhareddy Institute of Technology, Chowdavaram, Guntur (DT), Andhra Pradesh, India[1]

B.Tech Students, Department of Information Technology, Kallam Haranadhareddy Institute of Technolgy, Chowdavaram, Guntur(DT), Andhra Pradesh, India[2, 3, 4, 5]

ABSTRACT: The area of medical science has attracted great attention from researchers. Several causes for human early mortality have been identified by a decent number of investigators. The related literature has confirmed that diseases are caused by different reasons and one such cause is heart-based sicknesses. Many researchers proposed idiosyncratic methods to preserve human life and help health care experts to recognize, prevent and manage heart disease. Some of the convenient methodologies facilitate the expert's decision but every successful scheme has its own restrictions. The proposed approach robustly analyses an act of Decision Tree, Random Forest, XGBoost and Hybrid Model. After analysing the procedure, the intended method smartly builds. Initially, the intention is to select the most appropriate method and analysing the act of available schemes executed with different features for examining the statistics.

KEYWORDS: Machine learning, Classification Technique, Decision Tree, Random Forest Stacking Classifier, supervised machine learning.

I. INTRODUCTION

Some of the related works represent various convenient methods with the implication but none of the methods aid professionals under different characteristics. Therefore, the design and implementation of these methods pave the way for further research. Additionally, the presented work indicates that the utilization of the data mining method works better than other approaches. With a discussion of research objectives, motivation, and key findings this chapter describes the contribution towards the direction to improve the QoS of the system. Selection and formation are the most appropriate features instead of employing a complete list of features that are associated with the selected dataset.The data mining is the process of extracting unknown and predictive information from huge amount of data. It is an innovative tool with great potential to help companies and mainly focus on the most essential information in their data warehouses. Most commonly data mining is also known as Knowledge Discovery in Databases (KDD). KDD is the important process of identifying valid, new, potentially useful, and finally understandable patterns in data. Knowledge discovery process has iterative sequential steps of processes and data mining is one of the KDD processes.

Many researchers have been keenly interested for performing research work using ensemble learning techniques during the last decade. Significant improvement in performances have been noted using ensemble approaches by several authors. These ensemble approaches have expanded their scope in variety of areas including healthcare,finance, insurance, automobile, manufacturing, bioinformatics, aerospace and many more. An ensemble ofclassifiers is a combination of individual classifiers that helps in the classification of new test samples by joining and assessing the individual techniques separately in some manner. In supervised learning, ensemble learning has become one of the profound areas of study among machine learning researchers.Research shows that this composite model using ensemble classification has mostly outperform the capability of one single model. Numerous experimental study and analysis have depicted that the generalized error has been reduced by combining the outputs of multiple classifiers. The key objective of using ensemble approaches has to build a model of ensembles that incorporates a combination of diversified individual classifier techniques with good accuracy. In this work, we have focused our interest in the field of Bioinformatics due to its huge popularity in the research domain. The various ensemble learning methods used in this paper are described as follows. Bagging is one of the commonly used ensembles learning technique based on Bootstrap sampling technique introduced by Bremen .In this samples called bags are created in order to construct the individual classifier of the same algorithm or to manipulate the selection of training data. Herethe data point is selected

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 3, March 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1103077|

randomly along with the replacement strategy i.e. some data is taken from random sample and some is missed from the original dataset. The separate classifier is trained from each bag. Bagging combines all the classifiers constructed in previous phase and test sample is predicted by giving votes to the classifiers individually .Boosting is another kind of influential ensemble methodology proposed by Freund and Shapira based on the learning in sequence . Firstly, the learning is done on the overall complete data set and then it is done on the outcome data received from the last learning performance. The idea is to explicitly change the weights of each and every training datapoint. The obtained weights are increased formisclassified sample data points and similarly decreased for the data points that are classified in the correct manner. Stacking is the technique introduced by Wolpert that involves the formation of linear combinations of different predictors to give overall improved results. It is composed of two phases. In the first phase different models such as J48,Random Forest are learned on the basis of dataset and their output would help in generating new dataset. In the second phase the dataset is used with the learning algorithm to give the final output.Although plethora of research work has been done in the ensemble learning domain, few key related studies have been mentioned here. A comparative study was done using three ensemble learning techniques which were Additive Regression, Bagging and Stacking applied on six algorithms of Machine Learning in the research work . Their work was also experimentally analysed using WEKA tool. Another study was conducted to improve the performance of one of the key machine learning algorithms i.e. C4.5 decision tree. They have compared the following ensemble approaches: Bagging, Boosting and Randomization. The results obtained has signified if there is no or very less classification noise, then randomization is competitive with bagging however boosting took up the lead. And bagging outperforms boosting and randomization in the case of considerable noise. Authors of compared the following ensemble learning methods, Ada-Boost, Logit Boost, and Random forest for the classification of breast cancer metastasis. Their experimental results have shown that the above-mentioned ensemble learning approaches performs much better than the individual machine learning classification algorithms such as logistic regression and SVM. Anotherstudy in the context of EEG signal classification [4] is conducted where three ensemble learning methods Bagging, Boosting and Random subspace were evaluated using KNN and C4.5 as base classifiers. The results have indicated the significance of base classifiers on the ensemble techniques as they have a valuable impact.

II.RELATED WORK

The paper titled "Comparative Analysis of Classification Techniques in Data Mining Using Different Datasets" by Ritu Sharma, Mr. Shiv Kumar, and Mr. Rohit Maheshwari was published in the International Journal of Computer Science and Mobile Computing (IJCSMC) in December 2015.[1]The paper focuses on the comparative analysis of different classification techniques used in data mining using various datasets. The authors have discussed the fundamental concepts of data mining, classification, and various classification techniques, such as Decision Tree, Naive Bayes, Random Forest, and Support Vector Machine (SVM). The authors have used four datasets, namely Iris, Wine, Pima Indian Diabetes, and Breast Cancer Wisconsin, to compare the performance of these classification techniques. The performance of each technique is evaluated based on accuracy, precision, recall, and F1-score. The results indicate that SVM outperforms the other classification techniques in most cases. However, the authors suggest that the choice of the classification technique depends on the nature of the dataset and the problem being addressed.

[2]The paper titled "Prediction of Occupational Accidents Using Decision Tree Approach" by Sobhan Sarkar, Atul Patel, Sarthak Madaan, and JhareswarMaiti was presented at the IEEE Annual India Conference (INDICON) in 2016. The paper proposes a decision tree approach for predicting occupational accidents in industrial settings. The authors have collected data from a steel plant in India to build the prediction model. The dataset includes various factors that could lead to accidents, such as the age of the worker, work experience, safety training, work location, and type of work. The authors have used the C4.5 decision tree algorithm to build the prediction model. They have also applied different evaluation metrics such as accuracy, precision, recall, and F1-score to evaluate the performance of the model. The results indicate that the proposed decision tree approach is effective in predicting occupational accidents. The authors have compared the performance of their model with that of other classification techniques such as logistic regression, k-nearest neighbour (KNN), and support vector machine (SVM). The decision tree approach outperforms these techniques in terms of accuracy and other evaluation metrics. Overall, this paper presents an innovative approach to predicting occupational accidents using decision trees and provides insights into the factors that contribute to accidents in industrial settings.

III.SUMMARY

The authors have collected data from a steel plant in India to build the prediction model. The dataset includes various factors that could lead to accidents, such as the age of the worker, work experience, safety training, work location, and type of work. The authors have used the C4.5 decision tree algorithm to build the prediction model. They have also

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 3, March 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1103077|

applied different evaluation metrics such as accuracy, precision, recall, and F1-score to evaluate the performance of the model.

The paper titled "A Comparative Study of Ensemble Learning Methods for Classification in Bioinformatics" by Aayushi Verma and Shikha Mehta was presented at the IEEE 7th International Conference on Cloud Computing, Data Science & Engineering – Confluence, in 2017.[3] The paper focuses on the comparative study of ensemble learning methods for classification in bioinformatics. The authors have discussed the fundamental concepts of ensemble learning and various ensemble methods such as bagging, boosting, and random forest. They have also discussed the use of these techniques in bioinformatics. The authors have used four datasets related to bioinformatics, namely Leukemia, Colon Cancer, Lung Cancer, and Breast Cancer, to compare the performance of these ensemble learning methods. The performance of each method is evaluated based on accuracy, sensitivity, specificity, and F1-score. The results indicate that the random forest technique outperforms the other ensemble methods in most cases. However, the authors suggest that the choice of the ensemble method depends on the nature of the dataset and the problem being addressed. Overall, this paper provides a useful overview of ensemble learning methods in bioinformatics and offers insights into the comparative analysis of these methods using various datasets. The results of this study can be useful for researchers and practitioners working in the field of bioinformatics.

The paper titled "A Novel Paradigm of Melanoma Diagnosis Using Machine Learning and Information Theory" by K. C. Giri, M. Patel, A. Sinhal, and D. Gautam was presented at the 2019 International Conference on Advances in Computing and Communication Engineering (ICACCE) in Sathyamangalam, Tamil Nadu, India.[4]The paper proposes a novel paradigm for diagnosing melanoma, a type of skin cancer, using machine learning and information theory. The authors have used a dataset of skin lesion images to build the diagnosis model. The dataset includes various features of skin lesions, such as size, shape, colour, and texture.The authors have used a combination of feature selection and feature extraction techniques to extract the relevant features from the dataset. They have also used various machine learning techniques such as decision trees, support vector machines, and neural networks to build the diagnosis model. The performance of each technique is evaluated based on accuracy, sensitivity, specificity, and F1-score.The results indicate that the proposed paradigm is effective in diagnosing melanoma. The authors have compared the performance of their model with that of other machine learning techniques. The proposed model outperforms these techniques in terms of accuracy and other evaluation metrics.Overall, this paper presents an innovative approach to diagnosing melanoma using machine learning and information theory and provides insights into the features that contribute to the diagnosis of melanoma. The results of this study can be useful for researchers and practitioners working in the field of medical diagnosis.

IV.EXISTING METHOD

With the rise of machine learning, computer approaches have been separated into two categories: classical methods and machine learning methods. This section explains how sentimental analysis is identified and how machine learning methods outperform traditional approaches. In this project, the present approach has a particular flow, and conventional sentimental analysis is also applied for development. However, it takes a considerable amount of memory, and the outcome is not precise.

DISADVANTAGES

- Low Accuracy
- High complexity.
- Requires skilled persons.
- Highly inefficient.

V.PROPOSED SYSTEM

After reviewing all the existing techniques, some of the researchers signifying the various advantages of each suggested technique and elaborated several restraints that are still associated with obtainable methods and highly affect the working behaviour of the techniques. Among several associated issues, some of the key restraints such as inflexibility time consuming for building a model ,alternative parameters, and inaccurate verdicts.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 3, March 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1103077|

Block Diagram:

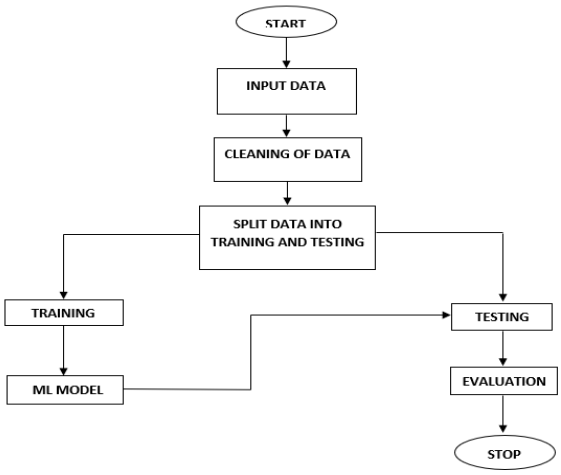


Fig.1 Block diagram of proposed method

ADVANTAGES

- Accuracy is good.
- Low complexity.
- Highly efficient.
- No need of skilled persons

VI.METHODOLOGY

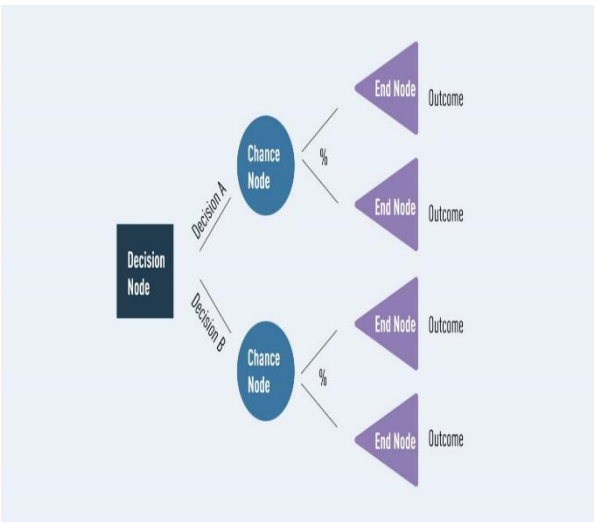
DECISION TREE

Decision tree is a type of flowchart that shows a clear pathway to a decision. In terms of data analytics, it is a type of algorithm that includes conditional 'control' statements to classify data. A decision tree starts at a single point (or 'node') which then branches (or 'splits') in two or more directions. Each branch offers different possible outcomes, incorporating a variety of decisions and chance events until a final outcome is achievedDecision trees are extremely useful for data analytics and machine learning because they break down complex data into more manageable parts. They're often used in these fields for prediction analysis, data classification, and regression

Decision trees can deal with complex data, which is part of what makes them useful. However, this doesn't mean that they are difficult to understand. At their core, all decision trees ultimately consist of just three key parts, or 'nodes':

- Decision nodes: Representing a decision (typically shown with a square)
- Chance nodes: Representing probability or uncertainty (typically denoted by a circle)
- End nodes: Representing an outcome (typically shown with a triangle)

Connecting these different nodes are what we call 'branches'. Nodes and branches can be used over and over again in any number of combinations to create trees of various complexity



Root nodes

In the diagram above, the blue decision node is what we call a 'root node.' This is always the first node in the path. It is the node from which all other decision, chance, and end nodes eventually branch.

Leaf nodes

In the diagram above, the lilac end nodes are what we call 'leaf nodes.' These show the end of a decision path (or outcome). You can always identify a leaf node because it doesn't split, or branch any further. Just like a real leaf!

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 3, March 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1103077|

Internal nodes

Between the root node and the leaf nodes, we can have any number of internal nodes. These can include decisions and chance nodes (for simplicity, this diagram only uses chance nodes). It's easy to identify an internal node—each one has branches of its own while also connecting to a previous node.

Splitting

Branching or 'splitting' is what we call it when any node divides into two or more sub-nodes. These sub-nodes can be another internal node, or they can lead to an outcome (a leaf/ end node.)

Pruning

Sometimes decision trees can grow quite complex. In these cases, they can end up giving too much weight to irrelevant data. To avoid this problem, we can remove certain nodes using a process known as 'pruning'. Pruning is exactly what it sounds like—if the tree grows branches we don't need, we simply cut them off.

Advantages of decision trees

- Good for interpreting data in a highly visual way.
- Good for handling a combination of numerical and non-numerical data.
- Easy to define rules, e.g. 'yes, no, if, then, else…'
- Requires minimal preparation or data cleaning before use.
- Great way to choose between best, worst, and likely case scenarios.
- Can be easily combined with other decision-making techniques.

Disadvantages of decision trees

- Over fitting (where a model interprets meaning from irrelevant data) can become a problem if a decision tree's design is too complex.
- They are not well-suited to continuous variables (i.e. variables which can have more than one value, or a spectrum of values).
- In predictive analysis, calculations can quickly grow cumbersome, especially when a decision path includes many chance variables.
- When using an imbalanced dataset (i.e. where one class of data dominates over another) it is easy for outcomes to be biased in favour of the dominant class.
- Generally, decision trees provide lower prediction accuracy compared to other predictive algorithms.

RANDOM FOREST:

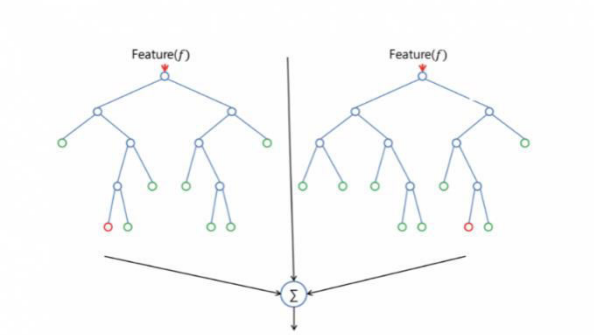
Random forest is a flexible, easy to use machine learning algorithm that produces, even without hyper-parameter tuning, a great result most of the time. It is also one of the most used algorithms, because of its simplicity and diversity (it can be used for both classification and regression tasks).Random forest is a supervised learning algorithm. The "forest" it builds, is an ensemble of decision trees, usually trained with the "bagging" method. The general idea of the bagging method is that a combination of learning models increases the overall result.

One big advantage of random forest is that it can be used for both classification and regression problems, which form the majority of current machine learning systems. Let's look at random forest in classification, since classification is sometimes considered the building block of machine learning. Below you can see how a random forest would look like with two trees:

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 3, March 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1103077|



Random forest has nearly the same hyperparameters as a decision tree or a bagging classifier. Fortunately, there's no need to combine a decision tree with a bagging classifier because you can easily use the classifier-class of random forest. With random forest, you can also deal with regression tasks by using the algorithm's regressor.
Random forest adds additional randomness to the model, while growing the trees. Instead of searching for the most important feature while splitting a node, it searches for the best feature among a random subset of features. This results in a wide diversity that generally results in a better model.
Therefore, in random forest, only a random subset of the features is taken into consideration by the algorithm for splitting a node. You can even make trees more random by additionally using random thresholds for each feature rather than searching for the best possible thresholds

IMPORTANCE:

The hyper parameters in random forest are either used to increase the predictive power of the model or to make the model faster

Increasing the predictive power
Firstly, there is the **nestimator's** hyper parameter, which is just the number of trees the algorithm builds before taking the maximum voting or taking the averages of predictions. In general, a higher number of trees increases the performance and makes the predictions more stable, but it also slows down the computation.
Another important hyper parameter is **max_features,** which is the maximum number of features random forest considers to split a node. SKlearn provides several options, all described in the documentation.
The last important hyper parameter is **min_sample_leaf.** This determines the minimum number of leaf's required to split an internal node.

2. Increasing the model's speed
The **jobs** hyper parameter tells the engine how many processors it is allowed to use. If it has a value of one, it can only use one processor. A value of "-1" means that there is no limit. The **random state** hyper parameter makes the model's output replicable. The model will always produce the same results when it has a definite value of random state and if it has been given the same hyper parameters and the same training data.

Advantages and Disadvantages of random forest

1) It can be used for both regression and classification problems.
2) Since base model is a tree, handling of missing values is easy.
3) It gives very accurate result with very low variance.
4) Results of a random forest are very hard to interpret in comparison with decision trees.
5) High computational time than other respective models.
Random Forest should be used where accuracy is up utmost priority and interpretability is not very important. Also, computational time is less expensive than the desired outcome.

XGBoost

XGBoost is the most popular machine learning algorithm.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 3, March 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1103077|

Extreme Gradient Boosting (XGBoost) is similar to gradient boosting framework but more efficient. It has both linear model solver and tree learning algorithms. So, what makes it fast is its capacity to do parallel computation on a single machine.
This makes XGBoost at least 10 times faster than existing gradient boosting implementations. It supports various objective functions, including regression, classification and ranking
XGBoost is an efficient and straightforward to use algorithm which delivers high performance and accuracy as compared to other algorithms. XGBoost is additionally referred to as regularized version of GBM. Let see a number of the benefits of XGBoost algorithm:

1. Regularization:XGBoost has in-built L1 (Lasso Regression) and L2 (Ridge Regression) regularization which prevents the model from over fitting. that's why, XGBoost is additionally called regularized sort of GBM (Gradient Boosting Machine) While using Scikit Learn library, we pass two hyper-parameters (alpha and lambda) to XGBoost associated with regularization. alpha is employed for L1 regularization and lambda is employed for L2 regularization.

2. Parallel Processing:XGBoost utilizes the facility of multiprocessing which is why it's much faster than GBM. It uses multiple CPU cores to execute the model. While using Scikit Learn library, nthread hyper-parameter is employed for multiprocessing. thread represents number of CPU cores to be used. If you would like to use all the available cores, don't mention any value for nthread and therefore the algorithm will detect automatically.

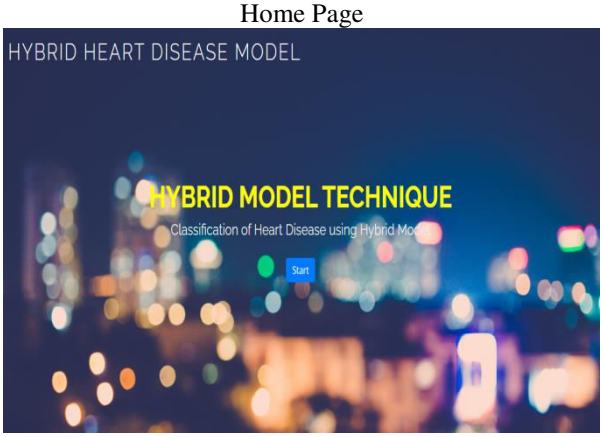
3. Handling Missing Values:XGBoost has an in-built capability to handle missing values. When XGBoost encounters a missing value at a node, it tries both the left and right split and learns the way resulting in higher loss for every node. It then does an equivalent when performing on the testing data.

4. Cross Validation:XGBoost allows user to run a cross-validation at each iteration of the boosting process and thus it's easy to urge the precise optimum number of boosting iterations during a single run. are often " this is often unlike GBM where we've to run a grid-search and only a limited values can be tested.

5. Effective Tree Pruning: A GBM would stop splitting a node when it encounters a negative loss within the split. Thus it's more of a greedy algorithm. XGBoost on the opposite hand make splits up to the max_depth specified then start pruning the tree backwards and take away splits beyond which there's no positive gain.

VI.RESULTS

Home Page

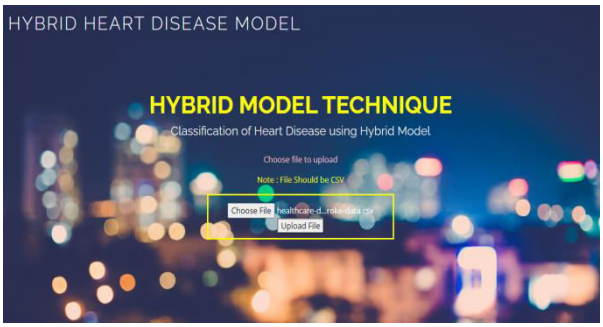


Upload Page

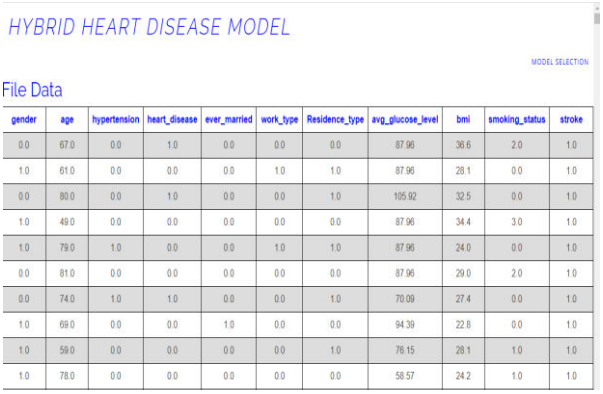
| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 3, March 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1103077|

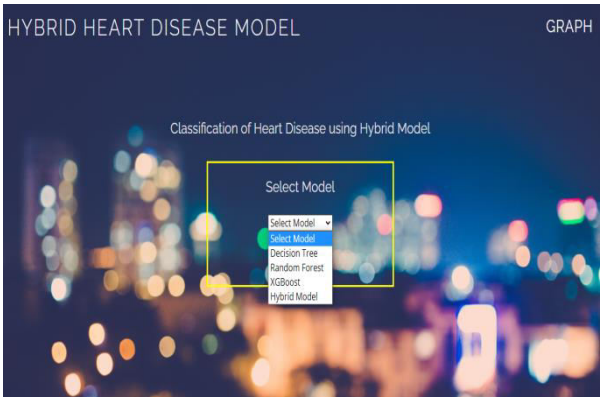


Data

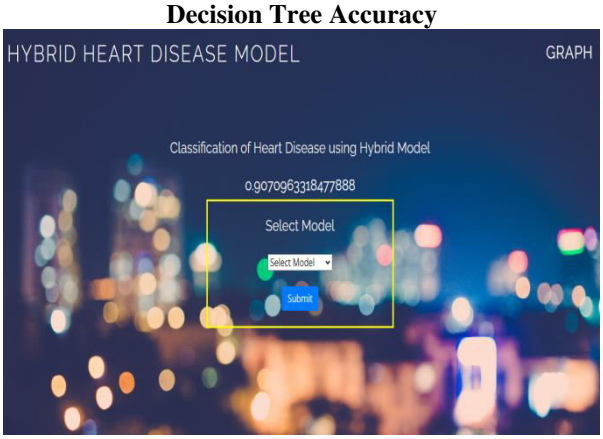


gender	age	hypertension	heart_disease	ever_married	work_type	Residence_type	avg_glucose_level	bmi	smoking_status	stroke
0.0	67.0	0.0	1.0	0.0	0.0	0.0	87.96	36.6	2.0	1.0
1.0	61.0	0.0	0.0	0.0	1.0	1.0	87.96	28.1	0.0	1.0
0.0	80.0	0.0	1.0	0.0	0.0	1.0	105.92	32.5	0.0	1.0
1.0	49.0	0.0	0.0	0.0	0.0	0.0	87.96	34.4	3.0	1.0
1.0	79.0	1.0	0.0	0.0	1.0	1.0	87.96	24.0	0.0	1.0
0.0	81.0	0.0	0.0	0.0	0.0	0.0	87.96	29.0	2.0	1.0
0.0	74.0	1.0	1.0	0.0	0.0	1.0	70.09	27.4	0.0	1.0
1.0	69.0	0.0	0.0	1.0	0.0	0.0	94.39	22.8	0.0	1.0
1.0	59.0	0.0	0.0	0.0	0.0	1.0	76.15	28.1	1.0	1.0
1.0	78.0	0.0	0.0	0.0	0.0	0.0	58.57	24.2	1.0	1.0

Model Selection



Decision Tree Accuracy

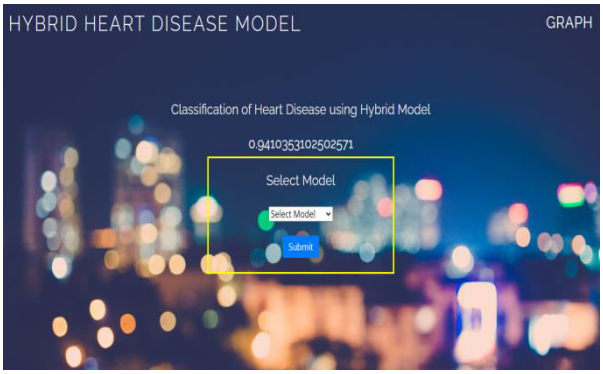


Random Forest Accuracy

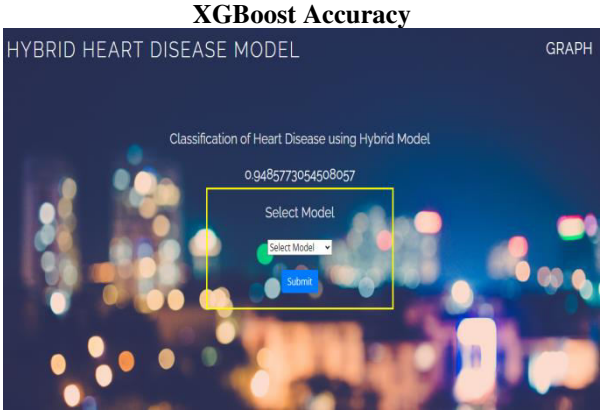
| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 3, March 2023 ||

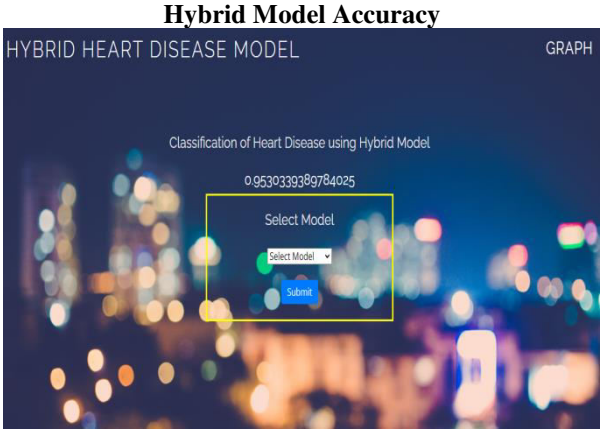
| DOI: 10.15680/IJIRCCE.2023.1103077|



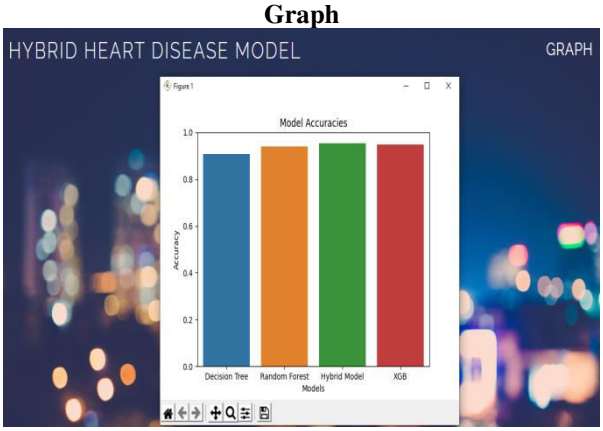
XGBoost Accuracy



Hybrid Model Accuracy



Graph

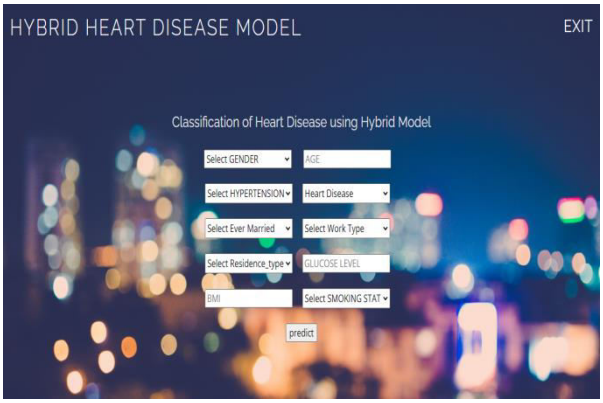


Prediction

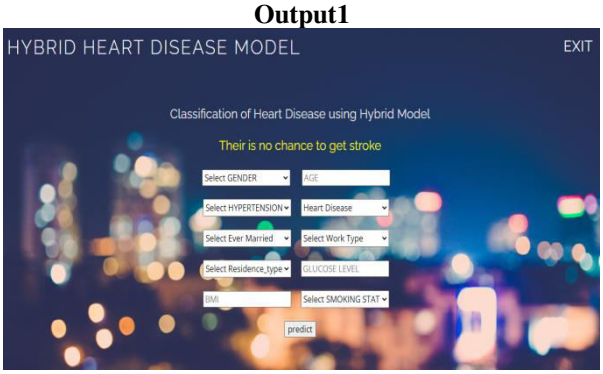
| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 3, March 2023 ||

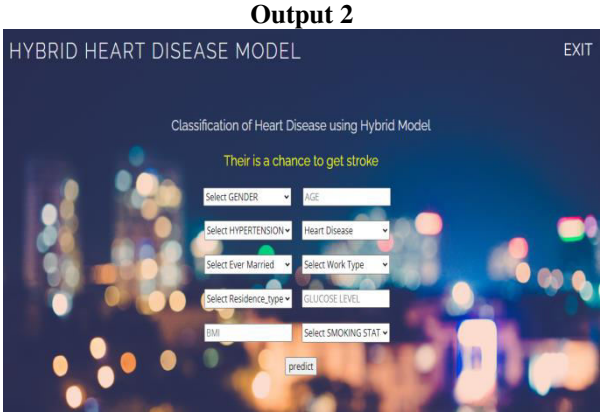
| DOI: 10.15680/IJIRCCE.2023.1103077|



Output1



Output 2



VII.CONCLUSION

The work that was carried out in this investigation endeavours to improve efficiency, suitability, and QoS. The characteristics and limitations of existing methods were discussed in the literature survey to build a more efficient method. The proposed work investigates four different algorithms such as the Random Forest, XGBoostand a form of Decision Tree (J48). The proposed method robustly analyses these four methods to exploited statistics and opts for the pair of the finest algorithm that utilizes a linear model based on the feature selection process with best-first search and Gain ratio along with the Ranker method. Several simulations have been carried out to demonstrate the efficiency of the proposed approach. Each comparison has indicated that the proposed approach effectively improves the issues of traditional as well as modern algorithms

REFERENCES

[1] Ritu. Sharma, Mr Shiv Kumar, Mr. RohitMaheshwari "Comparative Analysis of Classification Techniques in DataMining Using Different Datasets" International Journal of Computer Science and Mobile Computing, IJCSMC, Vol. 4, Issue. 12, December 2015, pp.-125 – 134.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 3, March 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1103077|

[2] SobhanSarkar, Atul Patel, SarthakMadaan, JhareswarMaiti "Prediction of Occupational Accidents Using DecisionTree Approach" IEEE Annual India Conference (INDICON), 2016, pp.- 1-6.
[3] AayushiVerma, Shikha Mehta "A Comparative Study of Ensemble LearningMethods for Classification in Bioinformatics" IEEE 7th International Conference on Cloud Computing, Data Science & Engineering – Confluence, 2017, pp.- 155-158.
[4] K. C. Giri, M. Patel, A. Sinhal and D. Gautam, "A Novel Paradigm of Melanoma Diagnosis Using Machine Learning and Information Theory," 2019 International Conference on Advances in Computing and Communication Engineering (ICACCE), Sathyamangalam, Tamil Nadu, India, 2019, pp. 1-7, doi: 10.1109/ICACCE46606.2019.9079975.
[5] AyisheshimAlmaw, KalyaniKadam "Survey Paper on Crime Prediction usingEnsembleApproach" International Journal of Pure and Applied Mathematics, Volume 118 No. 8 2018, pp.-133-139.
[6] ShakuntalaJatav and Vivek Sharma "An Algorithm For Predictive DataMining Approach In Medical Diagnosis" International Journal of Computer Science & Information Technology (IJCSIT) Vol 10, No 1, February 2018, pp.- 11-20.
[7] Han Wu, Shengqi Yang, Zhangqin Huang, Jian He, Xiaoyi Wang "Type 2 diabetes mellitus prediction model based on data mining" ELSEVIER Informatics in Medicine Unlocked, 2018, pp.- 100-107.
[8] Patel M., Choudhary N. (2017) Designing an Enhanced Simulation Module for Multimedia Transmission Over Wireless Standards. In: Modi N., Verma P., Trivedi B. (eds) Proceedings of International Conference on Communication and Networks. Advances in Intelligent Systems and Computing, vol 508. Springer, Singapore. https://doi.org/10.1007/978-981-10-2750-

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.379

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details