International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.025 ISSN: 2581-4621

Heart Disease Diagnosis and Prediction using Multi Linear Regression

Shah Alam[1], Bhaskar Bakshi[2], Rupjit Maity[3], Sulekha Das [4], Dr. Avijit Kumar Chaudhuri [5]
[1] UG-Computer Science and Engineering, Techno Engineering College Banipur,
[2]UG-Information Technology, Techno Engineering College Banipur,
[3]UG-Computer Science and Engineering, Techno Engineering College Banipur,
[4]Assistant Professor, Computer Science and Engineering, Techno Engineering College Banipur,
[5] Assistant Professor, Computer Science and Engineering, Techno Engineering College Banipur,
[1]Orcid Id : 0009-0009-6711-6247, [2]Orcid Id :0009-0003-3002-6324
[3]Orcid Id :0009-0002-9406-7291, [4]Orcid Id : 0000-0002-6641-3268
[5]Orcid Id : 0000-0002-5310-3180

ABSTRACT
The correct prediction of heart disease can prevent life threats, and incorrect prediction can prove to be fatal at the same time. In this paper machine learning algorithm is applied to compare the results and analysis of primary dataset. The dataset consists of 46 attributes among these Information gain is used to select 24 features for performing the analysis. Various promising results are achieved and are validated using accuracy and confusion matrix. The dataset consists of some irrelevant features which are handled and data are also normalized for getting better results. Using machine learning approach, 77.78% accuracy was obtained. Multiple linear regressions are used to construct and validate the prediction system. Our experimental result shows that multiple linear regressions are suitable for modelling and predicting cholesterol.
Keywords- Primary Dataset, Information Gain, 24 Attributes, Analysis, Multiple Linear Regression

INTRODUCTION

Heart attack diseases remains the main cause of death worldwide, including India and possible detection at an earlier stage will prevent the attacks. Medical practitioners generate data with a wealth of hidden information present, and it's not properly being used effectively for predictions. For this purpose, the research converts the unused data into a dataset for modelling using different data mining techniques. People die having experienced symptoms that were not taken into considerations. There is a need for medical practitioners to predict heart disease before they occur in their patients. The features that increase the possibility of heart attacks are smoking, lack of physical exercises, high blood pressure, high cholesterol, unhealthy diet, harmful use of alcohol, and high sugar levels. Cardio Vascular Disease (CVD) incorporates coronary heart, cerebrovascular (Stroke), hypertensive heart, congenital heart, peripheral artery, rheumatic heart, inflammatory heart disease. Data mining is a knowledge discovery technique to analyze data and encapsulate it into useful information. The current research intends to predict the probability of getting heart disease given patient data set. Predictions and descriptions are principal goals of data mining, in practice. Description emphasizes on discovering patterns that explains the data to be interpreted by humans. The purpose of predictions in data mining is to help discover trends in patient data in order to improve their health. Due to change in life styles in developing countries, like South Africa, Cardio Vascular Disease (CVD) has become a leading cause of deaths. CVD is projected to be a single largest killer worldwide accounting for all deaths. An endeavor to exploit knowledge, experience and clinical screening of patients to diagnose or recognize heart attacks is regarded as a treasured opportunity. In the health sectors data mining plays an important role to predict diseases. The predictive end of the research is a data mining model. In this paper, a machine learning method is applied to investigate information regarding Heart Diseases, to assess the prescient intensity of these systems. To this aim, Multiple Linear Regression algorithm is first developed to predict cholesterol and High Blood Pressure in its early stage. As the feature selection algorithm can affect the performance of the Multiple Linear Regression model, a

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.025 ISSN: 2581-4621

Information Gain is utilized to optimize the model used to predict. This enables the model to achieve better accuracy in the prediction and prognosis stages. Besides, the value of the coefficients β_0, β_1, β_2, ..., β_n in the Multiple Linear Regression algorithm is determined experimentally using an iterative approach. In the end, the performance of the proposed algorithm is assessed when it applies to a Heart-Disease database.

Reference	Method	Key Findings	Dataset	Challenges
[1]	Feature selection algorithm (FCMIM), SVM	Improved accuracy results for heart disease dataset	Heart disease dataset (Cleveland)	Perform better only for small dataset
[2]	Hybrid Machine Learning, Hybrid Random Forest	Better accuracy (87.8 %)	Cardiovascular disease dataset	Limited features
[3]	IoT, Machine Learning Methods, SVM	Accuracy (97.5 %)	Heart Disease dataset	Limited features
[4]	Various Machine Learning classification techniques and Principal component analysis have been used to anticipate heart disease	Better measurements and select characteristics results	Hungarian-Cleveland data	Dimension issues, accuracy
[5]	Naive Bayes and SVM were used as classifiers	Classification of heart disease dataset, cause of heart disease, diabetes	Kaggle dataset	Features selection and classification performs slower
[6]	k-NN algorithm	Feature selection, Classification	Kaggle dataset	Feature Categorization can be improved.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.025 ISSN: 2581-4621

Literature Review

Machine learning involves several algorithms such as Regression Analysis (Linear, Multiple, Logistic), decision trees, random forests, k-Nearest Neighbors (kNN), support vector machine (SVM), Naive Bayes (NBs), classification tree (C4.5), gradient boosting machines (GBM), etc. While each of these algorithms processes data differently, in this section, a few recently proposed machine learning candidates in the area of malignant growth finding are reviewed chronologically.

Below are some, Predictive Analysis of Heart Diseases with Machine Learning Approaches

Overall, these studies demonstrate the utility of various machine learning and statistical algorithms in investigating the complex relationships between risk factors and heart disease, and in predicting the risk of adverse outcomes. By incorporating these advanced analytical tools into clinical practice, researchers and clinicians can better identify and manage high-risk patients, ultimately leading to improved cardiovascular health outcomes.

The Proposed Approach

The proposed methodology is an enhancement of the Multiple Linear Regression Method for Test-Train-Partition and Random Forest for 10-Fold Cross Validation. This section briefly provides a background for the Multiple Linear Regression Method.

Multiple Linear Regression

Multiple linear regression is a statistical method used to model the relationship between a dependent variable and multiple independent variables. The basic idea is to use the values of the independent variables to predict the value of the dependent variable.

The multiple linear regression model can be represented mathematically as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_n X_n + \varepsilon$$

Where:

Y is the dependent variable or the response variable that we want to predict

$X_1, X_2, \ldots, X_n$ are the independent variables or predictors

$\beta_0, \beta_1, \beta_2, \ldots, \beta_n$ are the coefficients or the model parameters that determine the relationship between the independent variables and the dependent variable

ε is the error term, which represents the random variability that cannot be explained by the independent variables.

Random Forest

Random Forest is a machine learning algorithm that is commonly used for classification and regression tasks. It is an ensemble method that combines multiple decision trees and generates a more accurate and robust prediction.

The algorithm works by creating a large number of decision trees (also called "forest"), each of which is trained on a random subset of the features and a random sample of the training data. This process is called "bagging" (short for "bootstrap aggregating"). Each decision tree is constructed by recursively splitting the data based on the feature that provides the most information gain or the best split according to some criterion (e.g., Gini impurity or information gain). Once all the decision trees have been constructed, the prediction of the random forest is obtained by averaging the output of all the individual trees (in the case of regression) or by taking a majority vote (in the case of classification).

Users can use both parameters of Gini impurity by default and set their variance as a substitute for categorization. In regression, both of the parameters use mean square error to determine variance reduction. Variability reduction can also be calculated in Scikit-learn using mean absolute error [37].

$$Gini\ Impurity = 1 - Gini \quad (5.1)$$

$$Gini = P_1^2 + P_2^2 * P_3^2 \ldots\ldots + P_n^2 \quad (5.2)$$

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.025 ISSN: 2581-4621

The Equation (5.1) represents the Gini impurity formula. **Where** P1 ……P_n represents the probabilities of each possible class in solution space, *Gini* represents the purity, and *Gini Impurity* represents the impurity of a particular node. Here *Gini* works only for categorical targets.
Random Forest has several advantages over individual decision trees. First, it reduces overfitting, which is a common problem with decision trees, by combining the predictions of multiple trees. Second, it is relatively insensitive to the choice of hyperparameters, such as the maximum depth of the trees, because the ensemble approach smooths out the noise in each individual tree. Third, it can handle a large number of features and a large training set, making it suitable for high-dimensional datasets.
Random Forest is widely used in various applications, such as image classification, speech recognition, credit scoring, and drug discovery, to name a few. Its flexibility, accuracy, and ease of use make it one of the most popular machine learning algorithms.

Feature Selection

The first goal in the proposed feature selection method is to reach at least the same accuracy rate as the whole features provide. The second goal is to improve the accuracy rate. Here, not only gathering extensive information on the features costs too much in terms of both the time and money, but also extra information results in wastage of time in classifying and diagnosis. As such, it is better to reduce the dimension in terms of the number of features to get a better response and to find a better correlation between the features and the outcomes.
The ***Information Gain*** algorithm is a technique to select the best features.

Information Gain Algorithm

The information gain algorithm is a popular method used in decision tree learning and feature selection. It is used to determine the relevance of a feature or attribute in predicting a target variable in a dataset.
The basic idea of the information gain algorithm is to calculate the amount of information provided by each attribute in the dataset and then select the attribute that provides the most information about the target variable.
Here's how the algorithm works:
Calculate the entropy of the target variable. Entropy is a measure of the randomness or uncertainty in the target variable. It is calculated as:
$H(Y) = -\Sigma\ p(y)\ \log 2\ p$
where p(y) is the proportion of samples that belong to class y.
For each attribute in the dataset, calculate the entropy of the target variable after splitting the dataset based on the values of the attribute. This is called the conditional entropy and is calculated as:
$H(Y|X) = \Sigma\ p(x)\ H(Y|X=x)$
where p(x) is the proportion of samples with attribute value x and H(Y|X=x) is the entropy of the target variable for the samples with attribute value x.
Calculate the information gain for each attribute as the difference between the entropy of the target variable before and after splitting the dataset based on the attribute. This is calculated as:
$IG(X) = H(Y) - H(Y|X)$
Select the attribute with the highest information gain as the next node in the decision tree.
The idea behind the information gain algorithm is that the attribute with the highest information gain provides the most information about the target variable and is therefore the most useful attribute for predicting the target variable. By recursively applying this algorithm to the remaining attributes, a decision tree can be constructed that can be used for classification or regression tasks.
However, the information gain algorithm has some limitations. For example, it may suffer from overfitting if the number of attributes is large compared to the number of samples in the dataset. To overcome this problem, various techniques such as pruning and regularization can be used.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.025 ISSN: 2581-4621

METHODOLOGY

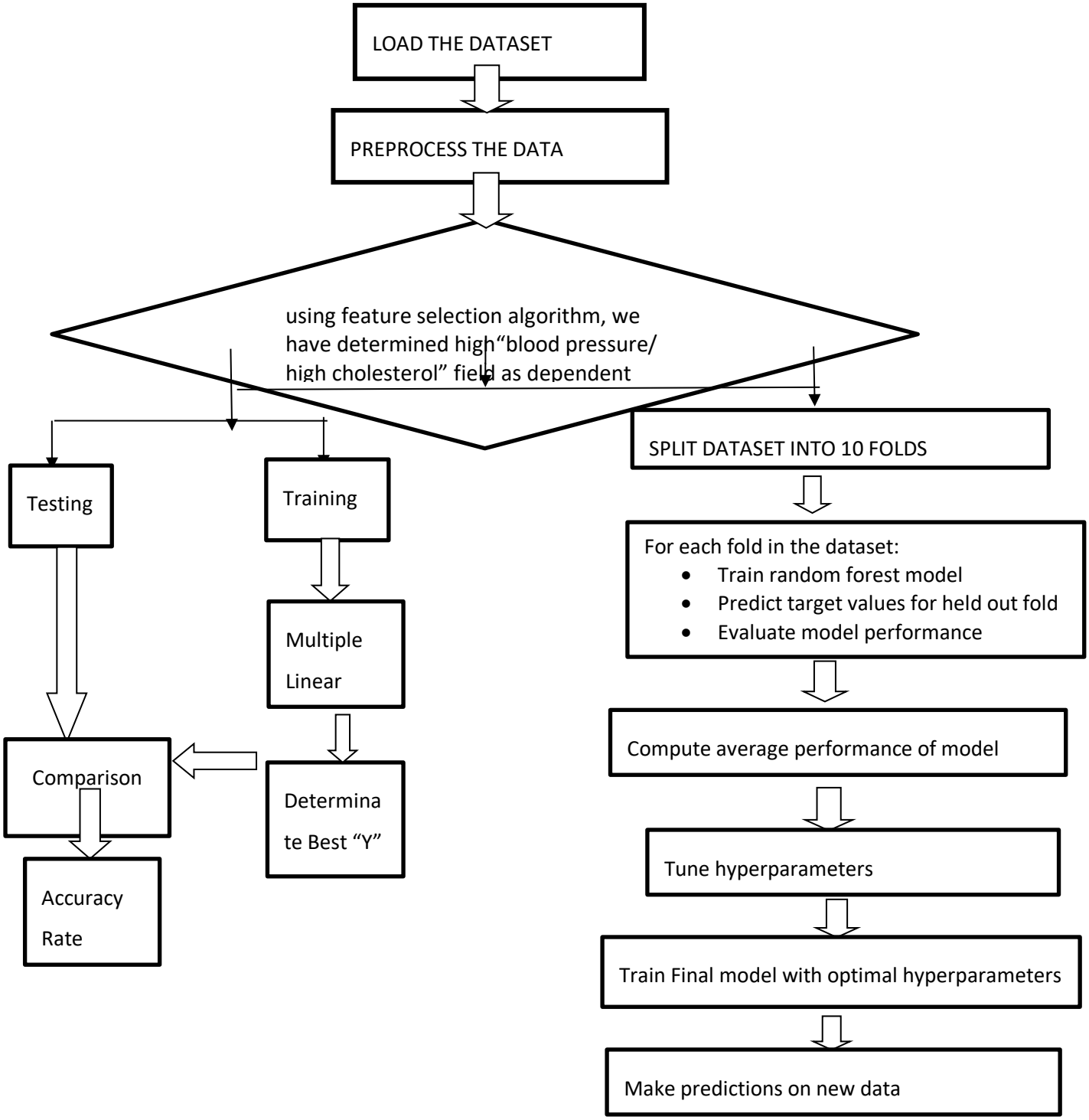
In this paper, we work in the following field…

Table.1.Data Field

Attributes	Description
Age	Age of a person
Mode of Transport	Type of vehicle
Do you take medicines for Diabetes?	Do the person medicine for diabetes?
Time to stay beyond office hours	How many hours the person stay in office?
Stops smoking recently?	Do the person quit smoking?
Are you trying to start walk or exercise?	Do the person exercise?
Do you take medicine for liver?	Do the person medicine for lever?
Food habit	Which type of food the person eat?
Industry	Type of the industry the person work at?
Do you play games?	Do the person play football, cricket etc.?
How many time you socialize through social media?	How many time the person active on Facebook, Instagram etc.?
How frequently you check mail or/and WhatsApp?	How many time he check mail and WhatsApp
Are you trying to give up taking meat?	Is the person trying to give up eating meat?
Do you frequently take egg?	How many times the person eat egg in a week?
Time to stay in office	How many hours the person work in the office?
Do you take medicine for high blood pressure?	Do the person medicine for high blood pressure?
Position in office	Name of the post the person work in office
Do you frequently take alcohol?	How many times the person drink alcohol in a week?
Like to drink coke or similar?	Is the person frequently drink coke?
Do you like to take salt additionally in food you eat?	Do the person take salt additionally in his food?
Do you like to take ice-cream?	Is the person frequently eat ice-cream?
Go for walk?	Do the person walk frequently?
Smoker?	Is the person smoker?
Like to eat pizza?	Is the person frequently eat pizza?

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.025 ISSN: 2581-4621

FLOW CHART



Using **Weka** (A popular open-source software tool for data mining and machine learning) , we processed the dataset for **Feature Selection** using **Information Gain** and selected the best features for maximum accuracy.

Then using **Multiple Linear Regression Algorithm** in case of train-test-partition of data and **Random Forest Algorithm** for 10-Fold cross validation.

Multiple linear regression is the strategy of statistics in regression that's familiar to analysing the link between one response variable (dependent variable) with 2 or additional controlled variables

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.025 ISSN: 2581-4621

(independent variables). This methodology was selected for this analysis as a result there have been quite controlled variables. during this analysis, the response variable is cholesterol in blood Only(Y). Age (X1), food habit (X2), do he smoke (X3) etc. are controlled variables.

- **Confusion-Matrix**

After finding the accuracy of the difference between actual data and calculated data we did the Confusion Matrix. In this confusion matrix it can be seen that,[2] we find the **TP** – which stands for **'TRUE POSITIVE'** means the accuracy of classified positive data, **TN** – which stands for **'TRUE NEGATIVE'** means the accuracy of classified negative data, **FP** – which stands for **'FALSE POSITIVE'**, means which remark that actual value is negative but predicted data is positive, **FN** – which stands for **'FALSE NEGATIVE'** means that actual data and the predicted data both are negative and append the TP, TN, FP, FN value in 2*2 matrix(mat1). After that, we find the accuracy, sensitivity, precision, recall, and specificity. This matrix contains all the raw information about the predictions done by a classification model on a given data set.[3]

- **Cross-Validation**

After finding the accuracy of the difference between actual data and calculated data we did cross-validation. In this cross-validation process first, we divide the whole list into 10 sub-list and then we find the accuracy of 10 sub-list elements we also find the Confusion Matrix of each Sub-list and we find the accuracy, and sensitivity, precision, recall, and specificity.

ACCURACY: It's the ratio of the correctly labeled subjects to the whole pool of subjects. Accuracy is intuitional.

PRECISION: Precision is the ratio of the correctly +ve labelled by our program to all +ve labeled.

RECALL: Recall means out of the total positive, what percentage are predicted positive.

SPECIFICITY: Specificity is calculated as the number of correct negative predictions divided by the total number of negatives.

- **ACCURACY= (TP+TN/ TP+TN+FP+FN) * 100**
- **PRECISION = (TP/FP+TP) *100**
- **RECALL= (TP/FN+TP) *100**
- **SPECIFICITY = (TN/TN+FP) * 100**

RESULT& DISCUSSION

Table.2.Accuracy of difference between Actual data and Calculated data

Accuracy of 90% Data as Training Data or (0.9)	**63.64**
Accuracy of 80% Data as Training Data or (0.8)	**69.23**
Accuracy of 75% Data as Training Data or (0.75)	**80.95**
Accuracy of 66% Data as Training Data or (0.66)	**76.67**

Table.3.Confusion Matrix & Corresponding Result

For 66% Training Data	**For50% Training Data**
Confusion Matrix: 22 7 **24 7** **Accuracy:**76.67	**Confusion Matrix: 29 8** **38 8**

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.025 ISSN: 2581-4621

Precision: 75.86 **Recall:**75.86 **Specificity:** 77.42	**Accuracy:**80.72 **Precision:**78.38 **Recall:**78.38 **Specificity:**82.61
Data For 80% Training	**Data For 90% Training**
Confusion Matrix: 11 6 **16 6** **Accuracy:**69.23 **Precision:** 64.71 **Recall:** 64.71 **Specificity:**72.73	**Confusion Matrix: 6 4** **8 4** **Accuracy:**63.64 **Precision: 60** **Recall: 60** **Specificity:** 66.67

Table.4. For 10-foldcross-validation Accuracy

TESTCASE	ACCURACYRATE(%)
1	71.43
2	71.43
3	85.71
4	78.57
5	61.54
6	84.62
7	84.62
8	69.23
9	53.85
10	76.92

Table.5. For 10-foldcross-validation Results

1st Fold	**2nd Fold**
Confusion Matrix:[[9 0 0 0] [2 1 0 0] [1 0 0 0] [1 0 0 0]] **Accuracy**:71.43 **Precision**:66 **Recall**:71	Confusion Matrix:[[9 0 0 0] [2 1 0 0] [1 0 0 0] [0 1 0 0]] **Accuracy**:71.43 **Precision**:59 **Recall**:71
3rd Fold	**4th Fold**
Confusion Matrix:[[9 0 0 0] [0 3 0 0] [1 0 0 0] [0 1 0 0]] **Accuracy**:85.71 **Precision**:74 **Recall**:86	**Confusion Matrix**:[[9 0 0 0] [1 2 0 0] [1 0 0 0] [0 1 0 0]] **Accuracy**:78.57 **Precision**:67 **Recall**:79

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.025 ISSN: 2581-4621

5th Fold	6th Fold
Confusion Matrix:[[8 1 0] [3 0 0] [1 0 0]] **Accuracy**:61.54 **Precision**:46 **Recall**:62	**Confusion Matrix**:[[9 0 0] [1 2 0] [0 1 0]] **Accuracy**:84.62 **Precision**:78 **Recall**:85

7th Fold	8th Fold
Confusion Matrix:[[9 0 0] [1 2 0] [1 0 0]] **Accuracy**: 84.62 **Precision**: 80 **Recall**: 85	**Confusion Matrix**:[[9 0 0] [3 0 0] [1 0 0]] **Accuracy**: 69.23 **Precision**: 48 **Recall**: 69
9th Fold	**10th Fold**
Confusion Matrix: [[7 2 0] [3 0 0] [1 0 0]] **Accuracy**: 53.85 **Precision**: 44 **Recall**: 54	**Confusion Matrix**:[[9 0 0] [2 1 0] [0 1 0]] **Accuracy**: 76.92 **Precision**:68 **Recall**: 77

CONCLUSION

While there are many machine-learning methods available in the literature whose performances depend on different aspects including the dataset they are applied on, in this paper, a machine-learning method called Multiple Linear Regression Method was hybridized with a feature-selection Information Gain algorithm to classify the patients having High blood pressure, High Cholesterol or having both or having none. The objective of using information gain algorithm was to determine the best combination of the features that minimize the overall miscalculation of the Multiple linear regression method. Moreover, the best value for the number of neighbors in the Multiple linear regression algorithm was determined using an algorithm coded in Python. It was shown that when the Multiple linear regression method is hybridized with a feature selection algorithm, the classification accuracy increases significantly. As it mentioned before, 24 features had been chosen via the Information Gain algorithm.

Future works may involve the use of other machine-learning classification algorithms or employing other population-based feature selection meta-heuristics and compare their performances to the one obtained by the proposed approach.

REFERENCES

[1] S. Sharma and M. Parmar, “Heart diseases prediction using deep learning neural network model,” Interna- tional Journal of Innovative Technology and Exploring Engineering (IJITEE), Vol. 9, No. 3, 2020, pp. 124– 137.

[2] K. Guleria, A. Sharma, U. K. Lilhore, and D. Prasad, "Breast Cancer Prediction and Classification Using Supervised Learning Techniques," Journal of Computational and Theoretical Nanoscience, Vol. 17, No. 6, 2020, pp. 2519–2522.
[3] A. N. Repaka, S. D. Ravikanti, and R. G. Franklin, "Design and implementing heart disease prediction using naives Bayesian," 2019 3rd International conference on trends in electronics and informatics (ICOEI), 2019, pp. 292–305.
[4] P. Singh, S. Singh, and G. S. Pandi-Jain, "Effective heart disease prediction system using data mining tech- niques," International journal of nanomedicine, Vol. 13, 2014, pp. 121–128.
[5] S. M. Nagarajan, V. Muthukumaran, R. Murugesan, R. B. Joseph, M. Meram, and A. Prathik, "Innovative feature selection and classification model for heart disease prediction," Journal of Reliable Intelligent Envi- ronments, 2021, pp. 1–11.
[6] A. Singh and R. Kumar, "Heart disease prediction using Machine Learning algorithms," 2020 international conference on electrical and electronics engineering (ICE3), 2020, pp. 452–457.
[7] D. Singh and J. S. Samagh, "A comprehensive review of heart disease prediction using Machine Learning," Journal of Critical Reviews, Vol. 7, No. 12, 2020, pp. 281–285.
[8] T. J. Peter and K. Somasundaram, "AN EMPIRICAL STUDY ON PREDICTION OF HEART DISEASE USING CLASSIFICATION DATA MINING TECHNIQUES," 2012.
[9] S. B. Patil and Y. S. Kumaraswamy, "Extraction of Significant Patterns from Heart Disease Warehouses for Heart Attack Prediction," International Journal of Computer Science and Network Security (IJCSNS), vol. 9, no. 2, pp. 228–235, 2009.
[10] K. Sudhakar, "Study of Heart Disease Prediction using Data Mining," vol. 4, no. 1, pp. 1157–1160, 2014.
[11] N. Pawar, U. K. Lilhore, and N. Agrawal, "A hybrid ACHBDF load balancing method for optimum resource utilization in cloud computing," International Journal of Scientific Research in Computer Science, Engineer- ing and Information Technology, Vol. 3307, 2017, pp. 367–373.
[12] Pal, S. S., Paul, S., Dey, R., Das, S., & Chaudhuri, A. K. Determining the probability of poverty levels of the Indigenous Americans and Black Americans in US using Multiple Regression.
7. Nusrath Unnisa A; Manjula Yerva; Kurian M Z. "Review on Intrusion Detection System (IDS) for Network Security using Machine Learning Algorithms". International Research Journal on Advanced Science Hub, 4, 03, 2022, 67-74. doi: 10.47392/irjash.2022.014
8. Mohammadibrahim Korti; Basavaraj S. Malapur; Smita Gour; Rajesh M. Biradar. "Shuchi 1.0: Robotic System For Automatic Segregation of Waste & Floor Cleaning". International Research Journal on Advanced Science Hub, 4, 02, 2022, 31-37. doi: 10.47392/irjash.2022.007
9. Bhuneshwari Nayak; Rachana Choudhary; Roymon M. G.. "Isolation, Screening and Morphological characterization of Laccase producing fungi". International Research Journal on Advanced Science Hub, 4, 02, 2022, 38-43. doi: 10.47392/irjash.2022.008
10. Khaled Salem Ahmad Amayreh; Ahmad Taufik Hidayah Bin Abdullah. "Conjunction in Expository Essay Writing by Jordanian Undergraduate Students Studying English as a Foreign Language (EFL)". International Research Journal on Advanced Science Hub, 4, 02, 2022, 24-30. doi: 10.47392/irjash.2022.006
11. Pravin T, M. Subramanian, R. Ranjith,Clarifying the phenomenon of Ultrasonic Assisted Electric discharge machining, "Journal of the Indian Chemical Society", Volume 99, Issue 10, 2022, 100705, ISSN 0019-4522, Doi: 10.1016/j.jics.2022.100705
12. R. Sudhakaran, P.S. Sivasakthivel, M. Subramanian, Investigations on the effect of surface coatings on the weldment properties on chromium manganese stainless steel gas tungsten arc welded plates, Materials Today: Proceedings,Volume 46, Part 17,2021, Pages 8554-8560, ISSN 2214-7853, Doi: 10.1016/j.matpr.2021.03.541
13. Sivalakshmi B; Mahisha sri K.B; Swetha P. "Survey on Mammogram, Ultrasound, MRI, Spectroscopy, Biopsy for Detecting Tumor in Breast". International Research Journal on Advanced Science Hub, 3, 2, 2021, 30-37. doi: 10.47392/irjash.2021.027

14. Hari Prasada Raju Kunadharaju; Sandhya N.; Raghav Mehra. "Detection of Brain Tumor Using Unsupervised Enhanced K-Means, PCA and Supervised SVM Machine Learning Algorithms". International Research Journal on Advanced Science Hub, 2, Special Issue ICSTM 12S, 2020, 62-67. doi: 10.47392/irjash.2020.262
15. Suresh P; Justin Jayaraj K; Aravintha Prasad VC; Abishek Velavan; Mr Gokulnath. "Deep Learning for Covid-19 Identification: A Comparative Analysis". International Research Journal on Advanced Science Hub, 4, 11, 2022, 272-280. doi: 10.47392/irjash.2022.068
16. Rajesh P.; Vetrivel Govindarasu. "Analyzing and Predicting Covid-19 Dataset in India using Data Mining with Regression Analysis". International Research Journal on Advanced Science Hub, 3, Special Issue 7S, 2021, 91-95. doi: 10.47392/irjash.2021.216
17. Menonjyoti Kalita; Golam Imran Hussain. "Determining the Influencing Factors of COVID 19 on Mental Health Using Neural Network". International Research Journal on Advanced Science Hub, 3, Special Issue 6S, 2021, 126-129. doi: 10.47392/irjash.2021.177
18. Ajitha K; Samuel Joseph C; Mahila Vasanthi Thangam D. "Online marketing of agricultural products during COVID pandemic: Farmers and customers perspectives". International Research Journal on Advanced Science Hub, 3, Special Issue 6S, 2021, 94-101. doi: 10.47392/irjash.2021.173
19. Yeshi Ngima; Dorjee Tsering. "Impact of COVID-19 on Education". International Research Journal on Advanced Science Hub, 2, Special Issue ICAMET 10S, 2020, 34-39. doi: 10.47392/irjash.2020.196
20. Siddavatam rammohan reddy; Balaji krushna potnuru. "3D Printing Innovation during Covid-19 Pandemic". International Research Journal on Advanced Science Hub, 2, 8, 2020, 62-67. doi: 10.47392/irjash.2020.95
21. Pooja Dahiya; Roopsi Kaushik; Anil Sindhu. "Corona virus: an Overview Along with Its Alternative Diagnostic Measures". International Research Journal on Advanced Science Hub, 2, Special Issue ICARD 2020, 2020, 163-169. doi: 10.47392/irjash.2020.113
22. Remya S. "Covid19 and Environment-A Theoretical Review from Higher Education Students Perspective". International Research Journal on Advanced Science Hub, 2, Special Issue ICARD 2020, 2020, 227-230. doi: 10.47392/irjash.2020.124
23. Mohd. Akbar; Prasadu Peddi; Balachandrudu K E. "Inauguration in Development for Data Deduplication Under Neural Network Circumstances". International Research Journal on Advanced Science Hub, 2, 6, 2020, 154-156. doi: 10.47392/irjash.2020.55
24. Salini Suresh; Suneetha V; Niharika Sinha; Sabyasachi Prusty; Sriranga H.A. "Machine Learning: An Intuitive Approach In Healthcare". International Research Journal on Advanced Science Hub, 2, 7, 2020, 67-74. doi: 10.47392/irjash.2020.67
25. Trupti S. Gaikwad; Snehal A. Jadhav; Ruta R. Vaidya; Snehal H. Kulkarni. "Machine learning amalgamation of Mathematics, Statistics and Electronics". International Research Journal on Advanced Science Hub, 2, 7, 2020, 100-108. doi: 10.47392/irjash.2020.72
26. Salini Suresh; Suneetha V; Niharika Sinha; Sabyasachi Prusty. "Latent Approach in Entertainment Industry Using Machine Learning". International Research Journal on Advanced Science Hub, 2, Special Issue ICARD 2020, 2020, 304-307. doi: 10.47392/irjash.2020.106
27. Suneetha V; Salini Suresh; Niharika Sinha; Sabyasachi Prusty; Syed Jamal J. "Enhancement in the World of Artificial Intelligence". International Research Journal on Advanced Science Hub, 2, Special Issue ICARD 2020, 2020, 276-280. doi: 10.47392/irjash.2020.132
28. Logeswari T.. "Performance Analysis of Ml Techniques for Spam Filtering". International Research Journal on Advanced Science Hub, 2, Special Issue ICIES 9S, 2020, 64-69. doi: 10.47392/irjash.2020.161
29. Hari Prasada Raju Kunadharaju; Sandhya N.; Raghav Mehra. "Detection of Brain Tumor Using Unsupervised Enhanced K-Means, PCA and Supervised SVM Machine Learning Algorithms". International Research Journal on Advanced Science Hub, 2, Special Issue ICSTM 12S, 2020, 62-67. doi: 10.47392/irjash.2020.262Export Citation

DOI:10.46647/ijetms.2023.v07i02.025 ISSN: 2581-4621

30. Maneesha M; Savitha V; Jeevika S; Nithiskumar G; Sangeetha K. "Deep Learning Approach For Intelligent Intrusion Detection System". International Research Journal on Advanced Science Hub, 3, Special Issue ICARD-2021 3S, 2021, 45-48. doi: 10.47392/irjash.2021.061
31. Kalki N; Karthick M; Mr Kavin; Keerthana S; Sangeetha K. "Advanced Face Mask Detection System". International Research Journal on Advanced Science Hub, 3, Special Issue ICARD-2021 3S, 2021, 112-115. doi: 10.47392/irjash.2021.076
32. Naveenkumar S; Kirubhakaran R; Jeeva G; Shobana M; Sangeetha K. "Smart Health Prediction Using Machine Learning". International Research Journal on Advanced Science Hub, 3, Special Issue ICARD-2021 3S, 2021, 124-128. doi: 10.47392/irjash.2021.079
33. Kavya Shakthi R.P; Kavin Raja A.S; Janani S.R; Sangeetha K. "Industrial Machine Identification Using Augmented Reality". International Research Journal on Advanced Science Hub, 3, Special Issue ICARD-2021 3S, 2021, 68-71. doi: 10.47392/irjash.2021.066
34. Salma Begum; Sampurna P.. "A Study on growth in Technology and Innovation across the globe in the Field of Education and Business". International Research Journal on Advanced Science Hub, 3, Special Issue 6S, 2021, 148-156. doi: 10.47392/irjash.2021.181
35. Dhanya S Karanth; Kumaraswamy H.V; Rajesh Kumar. "Workaround prediction of cloud alarms using machine learning". International Research Journal on Advanced Science Hub, 3, 8, 2021, 164-168. doi: 10.47392/irjash.2021.231
36. Gyanendra Kumar Pal; Sanjeev Gangwar. "Discovery of Approaches by Various Machine learning Ensemble Model and Features Selection Method in Critical Heart Disease Diagnosis". International Research Journal on Advanced Science Hub, 5, 01, 2022, 15-21. doi: 10.47392/irjash.2023.003