International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 05 July 2023 **Revised:** 25 August 2023 **Accepted:** 15 September 2023*

Novel Classification and Prediction of Heart Disease using CDMA Algorithm

Tintu George[1], A.Hema[2]
[1]Research Scholar
Department of Computer Science
Kongunadu Arts and Science College, Coimbatore, India
tintueldo@gmail.com
[2]Associate Professor & Head
Department of Information Technology
Sankara College of Arts and Commerce, Coimbatore, India
hemasrgm@yahoo.com

Abstract— Chronic illness is a long-term condition that lasts a lifetime. In most cases, immunizations and medications cannot heal them, or they do not work. The most common chronic illnesses are heart disease. The first step in stopping the progression of these disorders is patient diagnosis and prognosis. The identification of individuals with heart disease may be made easier with the machine learning (ML) and deep learning (DL). Finding people who is at risk for these well-known illnesses is often influenced by a variety of circumstances. High precision is provided by deep learning. Machine learning, however, provides less precision. Deep learning also needs a lot of data. However, machine learning can be trained on less data. By doing so, we may determine that one technique's flaw is fixed by another. To classify and forecast heart disease, this research developed an algorithm by combining ML and DL algorithm that is Combination of Machine Learning and Deep Learning Algorithm (CMDA). The data set for the work was taken from UCI data repository. The CMDA algorithm uses the Dl4jMlpClassifier and the Support Vector Machine (SVM). The technique like stacking classifier is used to integrate above two algorithms in the CMDA. The classification method utilized Naive Bayes as a meta-classifier in the CMDA algorithm, uses a stacking classifier strategy for final prediction. After prediction finally, the CMDA method utilizes the Min-Max normalization approach to determine risk factor. According to the experimental findings, the proposed CMDA algorithm effectively classifies and forecasts heart disease and produce high results while comparing with existing methods.

Keywords- Classification, prediction, stacking, chronic disease, normalization, machine learning, deep learning

I. INTRODUCTION

Chronic disease is a long-lasting illness that persists over period of time. The health of the disabled population may also be impacted by chronic illness since it may impose extra activity limitations [1]. Even if they don't have any symptoms, people with chronic diseases usually think they are well. The lack of any symptoms, however, does not imply that a chronic illness has vanished [3]. In general, the most common chronic illnesses are heart disease. Early illness detection helps to reduce medical costs and the risk of people developing highly complicated health problems [20].

In machine learning, it extracts the feature and deliver the samples for classification or prediction. Thus, these two stages take place on their own [2]. However, it is difficult to accurately extract characteristics from vast amounts of data and then send them to a sample for categorization. As a result, the DL approach benefits from the massive compute energy that is now available as well as the enormous raw data that is further implemented across several deep layers[22]. The output layer, which is also the layer that finally produces the classification results. As a result, it observes that both feature extraction and categorization take place inside the deep layers [5].

The detection heart disease with people is suggested to be done using machine learning and deep learning techniques [21]. Finding people who are at risk for these prevalent illnesses depends on a variety of criteria. To address difficulties with heart disease, ML and DL approaches are taken into consideration with feature extraction, clustering, classification, and prediction.

The diagnosis of heart ailment can be discovered using classification technique like twin classification. The twin classification has 2 sets as heart disease is present or no heart disease not present [19]. This classification algorithm consists of two type of data namely training data and testing data. The training data is used to grasp the input data that are consociated with the type. Besides, this algorithm is conditioned exactly, and it can be applied to find whether a particular person has heart disease or not.

As a result, this novel approach for heart disease classification and prediction utilizing ML and DL algorithms (CMDA) is provided. The CMDA algorithm uses the Dl4jMlpClassifier method for DL and the Support Vector Machine (SVM) algorithm is considered for ML. The stacking classifier is used to integrate two algorithms in the CMDA. The Naive Bayes algorithm used for classification it is utilized as meta-classifier in CMDA and stacking classifier for final prediction work. Finally, the CMDA method utilizes the Min-Max normalization approach to determine risk factor.

II. RELATED WORK

A. Classification

Using ML algorithm akin to RF (Random Forest algorithm) and feature selection methods like entropy evaluation technique-based forward selection with backward elimination technique by using percentage split like test choice, Raghavendra et al. [6] evaluated the PIMA Diabetes data set. The R studio platform was used to carry out the experiment. The authors also achieved an 84.1% classification accuracy. The scientists inferred from the results that RF predicted diabetes better than other technologies by a lower number of characteristics, allowing one to avoid even the smallest important test for detecting diabetes.

A modified calculation for classification using decision trees was described by Mathan et al. [10] and provides accurate results when compared to existing methods. They use the greedy approach to deal with the most excellent aspects in their work to display the data mining strategy at illness prediction frameworks in the medical sector. Their research demonstrates that among the many forecasting methods, neural networks and Gini index forecasting methods provide the most amazing accuracy for heart attack prediction.

It is acknowledged that a discretization strategy comparable to the voting approach may carry more precise decision trees. Their study work monitors the outcome after the application of variety procedures. To improve implementation at coronary disease detection, different decision algorithm is used. The performance measures like accuracy and sensitivity are attained to the use of decision tree approaches.

Huang started to increase the classification accuracy when Chintan et al. [11] reported an efficient cardiac disease prediction utilizing machine learning approaches employing clustering. Models like the multilayer perceptron (MP), XGBoost (XGB), random forest (RF), decision tree classifier, are employed. According to author's analysis, the research's accuracy was 87.28% and was based on multilayer perceptron's with cross-validation and other techniques.

B. Prediction

An accurate technique for diabetes prediction was reported by Kadhm et al. [7]. This method reduced execution time by deleting unnecessary information using the K-nearest neighbor (KNN) algorithm. However, their categorization strategies assign each data model to the appropriate class utilizing a Decision Tree (DT). Through testing, their solution outperformed the prior method while using the Pima Indians Diabetes (PID) dataset for categorization.

A deep neural network was used by Ayon et al. [8] to predict the diabetes disease. This deep neural network technology was used in this research to identify diabetes utilizing a variety of medical parameters. The accuracy was done using five-fold cross-validation, which is somewhat more effective than other previously used methods to anticipate diabetes mellitus, is found to be higher by the authors. Their plan would benefit not just regular people but also medical staff.

Mohan et al. [9] introduced a method to improve the accuracy and prediction of cardiovascular illness, that aims to find the significant features selection by using Machine Learning technologies. The forecasting method is based on multiple well-known classification technologies and various feature mixes. The mixed random forest algorithm with a linear prediction model (HRFLM) approach was used by the authors to enhance the efficiency by the performance measure of accuracy 88.7% for heart disease.

By using multiple kinds of data mining methods such as Decision Trees, Naive Bayes and Forest Logistic Regression. Apurb et al. [12] have suggested the study to predict Heart Disease and classify the patient's risk. And by examining the effectiveness of machine learning algorithms, the author gives a comparison of various methods. The Random Forest algorithm has a 90.16% accuracy rate.

Based on medical characteristics, Harshit et al. [13] have concentrated on cardiac illness. To predict and categorize cardiac disease, logistic regression and KNN are utilized, and they are compared to earlier classifiers like Naive Bayes.

A machine learning technique has been presented by Arslan Khan et al. [14] to predict the precision and decision-making of disease in patients for CVD. SRM was used to select the patients from both Pakistan's Lady Reading Hospital and Khyber Teaching Hospital. Patients with CVD were classified and predicted with the help of ML

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 05 July 2023 **Revised:** 25 August 2023 **Accepted:** 15 September 2023*

techniques such as DT, RF, LR, NB, and SVM. With the help of confusion matrix and the recursion operating curve characteristic the forecast were estimated for all the algorithm. The effectiveness of the ML algorithm was calculated under various scenarios.

III. DATASET DESCRIPTION

The proposed study was carried out with, the 270 rows and 14 columns of data on heart illness which were taken from the UCI machine learning library.

IV. PROPOSED MODELLING

Among the Machine Learning techniques, big data techniques particularly many features selections is arising these days[18]. Multiple investigators concentrate on experimentations to work these fates to passage necessary features from these important data. Statistical approaches were applied to remove the noise and unrelated data from the data set. To season the algorithm, we have not used with any features.

Our algorithm can be raised with features applicable and unrelated; therefore, the selection of features plays an essential part in clustering and classification. The selection of Feature means the system of relating and deleting inapplicable features from the data set. It reduces the bulk of the data and allow the learning methods to serve fast and more efficiently this helps to improve the performance measures.

Thus, a complete feature selection approach is necessitated. To find the problem, the clustering and classification algorithms are necessary. After clustering technique, the classification is done. In clustering the similar data are clustered and based on that the classification is done to enhance the performance measures.

A. Feature Selection

In feature selection, three alternative methods are used to choose the characteristics. Three distinct kinds of algorithms have been used in a three-tier feature selection approach. The feature is selected and extracted in tier 1's information gain-based filter.

Tier 2 employs the J48-based Wrapper feature approach. The Tier 3 data was subjected to the RIDGE Filter-based in Tier 1, wrapper-based in Tier 2, and embedded-based in Tier 3. The data is sent to the three tiers as input after pre-processing. Regression-based embedded feature selection approach. The suggested three-tier feature selection is shown in Figure 1 by using the best features from filter-based approach, wrapper-based approach, and embedded-based systems [15]. The best characteristics are retrieved and shown in Table 1 after feature selection.

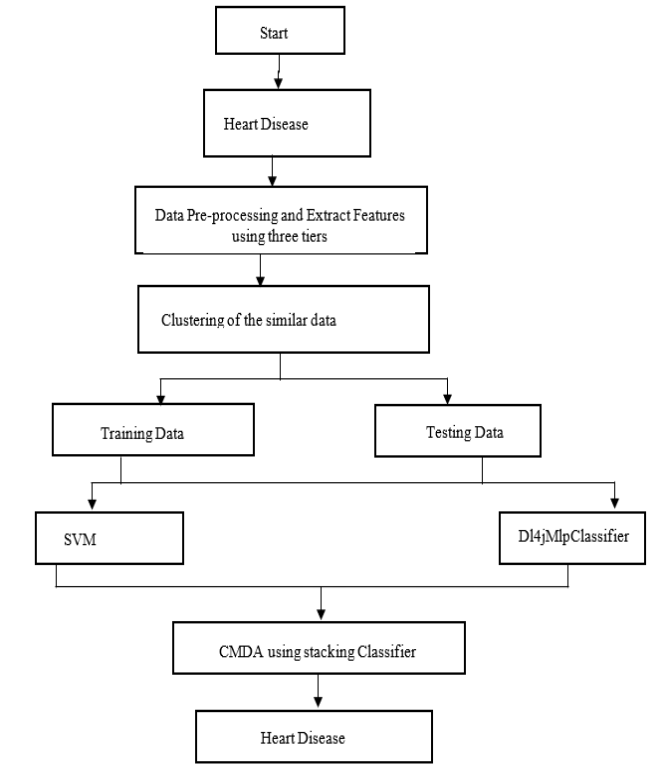

Figure 1: Overall flow of heart disease prediction

B. Feature Selection

In feature selection, three alternative methods are used to choose the characteristics. Three distinct kinds of algorithms have been used in a three-tier feature selection approach. The feature is selected and extracted in tier 1's information gain-based filter.

Tier 2 employs the J48-based Wrapper feature approach. The Tier 3 data was subjected to the RIDGE Filter-based in Tier 1, wrapper-based in Tier 2, and embedded-based in Tier 3. The data is sent to the three tiers as input after pre-processing. Regression-based embedded feature selection approach. The suggested three-tier feature selection is shown in Figure 1 by using the best features from filter-based approach, wrapper-based approach, and embedded-based systems [15]. The best characteristics are retrieved and shown in Table 1 after feature selection.

Table 1: Heart Disease Data Feature Information

Id	**Feature Type**	**Measues**
F1	chest pain type	1, 2, 3, 4
F2	maximum heart rate achieved	Numerical
F3	exercise-induced angina	0, 1
F4	old peak	Numerical

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *05 July 2023* ***Revised:*** *25 August 2023* ***Accepted:*** *15 September 2023*

F5	number of major vessels coloured by fluoroscopy	0, 1, 2, 3
F6	Thal	3, 6, 7
F7	Class	1 (Absence), 2(Presence)

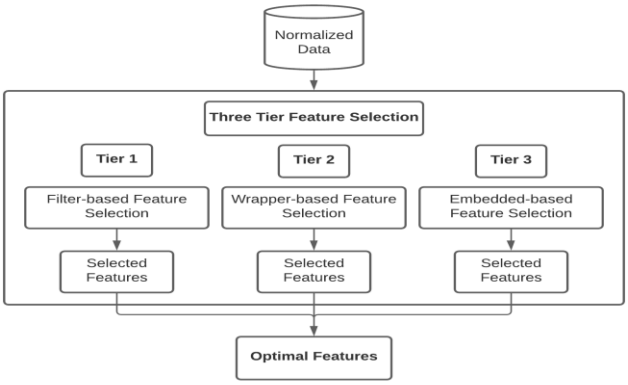

Figure 2: Three Tier Feature Selection [15]

Figure 2 shows the overall structure of three tier feature selection and Figure 2 displays the results of the heart disease dataset's in three Tier's of feature selection.

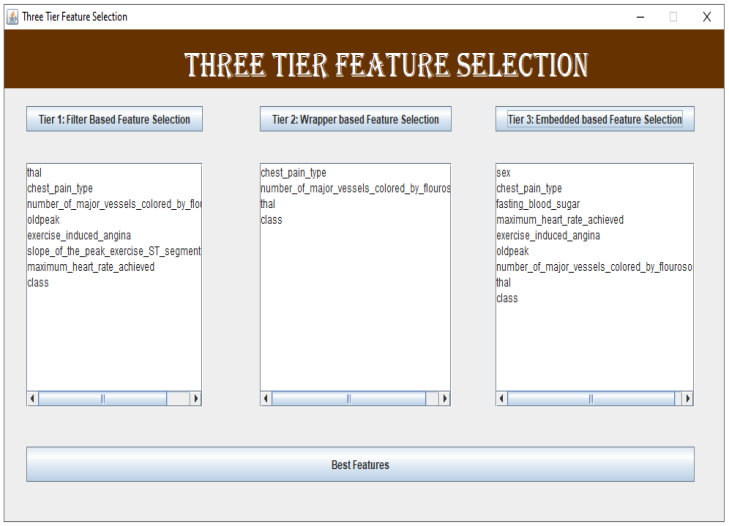

Figure 3: Tier 1, Tier 2 and Tier 3 feature selection outcomes for heart disease dataset [15]

The Heart disease dataset's common "true" condition fulfilment for the feature subset was chosen based on Figure 3.

C. Clustering

After the best feature is selected the using three tier feature selection, the output of feature selection is sent as the input to clustering. In clustering the similar data are grouped together by using four types of clustering algorithms. Partition clustering method, Hierarchical clustering method, Density based clustering method and distribution-based clustering method by using these four methods a new novel clustering framework was developed.

Input:	UCI machine learning library
Output:	Assigning each Data Instance (DI) to Majority Voted Technique (MVT)
Step 1:	Load the selected feature of UCI Dataset
Step 2:	Applying K-Means Clustering technique for UCI dataset
Step 3:	Applying Hierarchical Clustering technique for UCI dataset
Step 4:	Applying Density based Clustering method for UCI dataset
Step 5:	Applying EM Clustering for UCI
Step 6:	For each DI in UCI dataset
Step 7:	Result 1 = Get result for UCI by K-Means cluster algorithm
Step 8:	Result 2 = Get result for UCI by the Hierarchical cluster
Step 9:	Result 3 = Get result for UCI by the Density-based cluster algorithm
Step 10:	Result 4 = Get the EM cluster result for UCI
Step 11:	If (Result 1 is equals to Result 2), then
Step 12:	MVT = Result 2
Step 13:	else if (Result 1 is equal to Result 3), then
Step 14:	MVT = Result 3
Step 15:	Else If (Result 1 is equal to Result 4), Then
Step 16:	MVT = Result 4
Step 17:	End if
Step 18:	End for

Algorithm 1: Clustering Algorithm

Partitioning Clustering

Partitioning Clustering method is a kind of clustering that partition the data's into non-hierarchical blocks. It is labelled the centroid predicated clustering methods. The most special algorithm used for partitioning clustering is K- means Clustering algorithm. In this algorithm, the dataset partitioned into a set of k collections, where K applied to delineate the number of groups that is pre-defined. The centre of the cluster crystallized in such an approach that the distance among the data points of one group is lowest as analogized to another cluster centroid. The main of this K-Means clustering algorithm is to reduce

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 05 July 2023 **Revised:** 25 August 2023 **Accepted:** 15 September 2023*

the squared error function and also the intra cluster difference which is showed in Eq. (1):

$$P = \sum_{j=1}^{k} \sum_{i=1}^{n} |xij - cj|^2 \quad \ldots\ldots (1)$$

Here P indicated the objective function,

n shows the instances count,

k indicates the clusters count and

$|xij - cj|^2$ calculates the Euclidean distance.

Hierarchical clustering

Hierarchical clustering could be applied as a replacement for partitioned clustering as there's no necessity of pre-specifying the number of clusters to crystallize. In this clustering, the dataset is breaking up into groups to construct a tree- alike structure, beyond known as a dendrogram. The compliances or any number of clusters could cherry-pick by gashing the tree at the right degree. arrangement for hierarchical clustering generally falls into two kinds.

• Divisive It's a "top down" approach every abidance begin in one cluster, also divides are enforced recursively as one move down the ordering.

• Agglomerative It's a "bottom up" approach all abidance initiates in its cluster, also clusters couples are combined like one moves up the ordering.

Density- Based Clustering

The density- based clustering type links the largely dense regions into clusters; the arbitrarily accommodated distributions created as lon this clustering algorithm, the points are distributed as core points, accessible points and outliers, as follows. The point " y " is accessible from " x " if there's a approach x1,., xn if x1 = x and xn = y, then each xi is 1 which can be straightly access from xi. A point "x" which is inside the core point if the minimum minPts points are within length ε(then ε is the consummate direction of the neighbourhood from x). Those points are called as straightly accessible from x. By description, there is no points that are straightly accessible and controlled from the non-core point.

Distribution Model- Based Clustering

Distribution model- based clustering, the dataset decoupled using the probability of how a data belongs to a distinct distribution. The

(i) E-step: Calculate the variable that is missing in dataset.

(ii) M-step: It will maximize the parameter which is present in the model dataset.

Majority Voting Technique

At final the voting clustering technique is used. A majority voting is added and more than the half of the votes emit. The arrangement behind the majority vote is that the conclusion of a commission is advanced than the award of individualities. In voting- based clustering method means that each data case in a dataset vote for the specific cluster in which it belongs and for its original collection in each other clustering outgrowth. The highest of these valuations denotes the stylish group for the data case. This clustering means that each of the data case should be clustered along with the conviction of the majority of the algorithms.

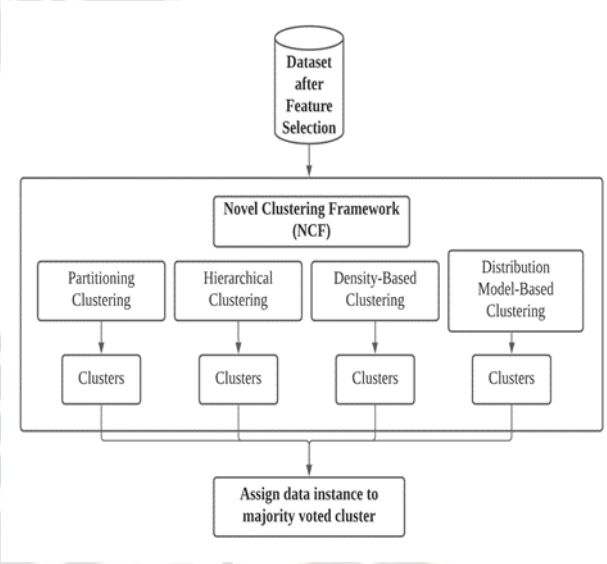

Figure 4: Novel Clustering Framework using Majority voting technique

D. Classification

Combination of ML and DL Algorithm:

Classification is the arrangement of effects or organisms into orderly collection using their parallel[17]. Classification is further called taxonomy. The DL approach is gathering actually celebrity because of its pre-eminence in accurateness when trained by the enormous amount of data. DL provides high accurateness. But ML gives lower accurateness [22]. Also, DL requires large data. But ML can train on lower data. DL takes further time for training. But ML takes lower time. Another approach solves the disadvantage of one fashion [23]. thus, this paper proposed a combination of ML and DL algorithm (CMDA) for heart disease classification and prediction. Support Vector Machine (SVM) algorithm used for ML and Dl4jMlpClassifier algorithm used for DL in the CMDA algorithm.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *05 July 2023* ***Revised:*** *25 August 2023* ***Accepted:*** *15 September 2023*

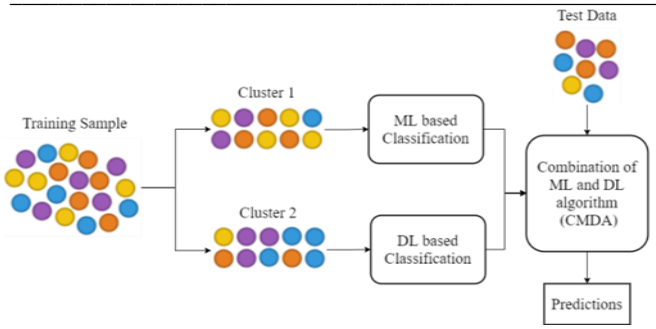

Figure 5: Flow diagram of CMDA

Figure 5 shows the inflow illustration of the CMDA algorithm. After clustering a training dataset using a novel clustering frame (NCF), ML- based classification applied for cluster 1 and DL based classification applied for cluster 2.

Input	UCI Training Dataset Cluster 1 (UTDC1), UCI Training Cluster 2 (UTDC2), UCI Testing Dataset (UTD)
Output	UCI Predicted Result (UPR)
Step 1:	Classify UTDC1 based on ML classifier SVM using weka
Step 2:	Classify UTDC2 based on Dl4jMlpclassifier using weka tool
Step 3:	For each instance of data (DI) from UTD
Step 4:	P1 = Data Instance Prediction using SVM classifier
Step 5:	P2 = Prediction of Data Instance using Dl4jMlpclassifier
Step 6:	UTD = Prediction of Data Instance using Stacking classifier (SVM + Dl4jMlpclassifier as stacking classifier and Naive Bayes classifier as Meta- Classifier)
Step 7:	End for

Algorithm 2: Combination of Deep Learning and Machine Learning Algorithm (CDMA)

This algorithm 2 shows Combination of ML and DL. In the CMDA algorithm, the stacking classifier approach is used to combine multiple classifier using meta classifier. Therefore, a new a new algorithm is proposed for stack classification and prediction called as combination of machine learning and deep learning algorithm (CMDA) for heart disease prediction. In CMDA, we employed the stacking algorithm which is used to classify the meta feature to produce the final class. In first layer (SVM and Dl4jMlpclassifier) which returns the probability. In the second layer the meta features are given as the input to the meta classifier (Naive Bayes classifier). Finally, the output of the classifier is produced.

Support Vector Machine

The SVM technique is a supervised ML technique which is using statistical proposition. SVM unnaturally works as the direct partition among two data points to honour two colorful classes at the multidimensional surroundings. The SVM algorithm unnaturally builds hyper plane surroundings therefore that each element is varied to the divided direct line. The hyper plane idea is handed to do the partition of the data using the large distance study to produce the class. To reduce the error rate of the algorithm, the biggest periphery classifier technique (SVM) is used.

Dl4jMlpClassifier

One of the deep learning classification algorithm is Dl4jMlpClassifier which allows the confirmation of arbitrarily deep feed forward neural network, it also include the concept of convolutional neural networks. Dl4jMlpClassifier technique is a core method of WekaDeeplearning4j, erected as a Weka package and thus creates Deeplearning4j styles accessible at the entire Weka terrain. The Dl4jMlpClassifier could be employed for together retrogression and bracket through opting suitable loss functions.

Stacking classifier approach

A mounding classifier follows a fashion and the affair from colorful of the classifier technique is transferred as input to the meta- classifier and the final result of classification is produced. Stacking classifier fashion could be an extremely competent system to execute a multi-classification issue. The entire individual bracket ways, generally known as base literacy ways, could be intermingled through erecting the meta classifier for ending results cast task. This could be completed by using mounding the results concertedly from each bracket algorithm and passing it as an input to the meta classifier. The Naïve Bayes classifier, is employed as meta- classifier in CMDA algorithm.

Naive Bayes classifier

Naive Bayes algorithm is a supervised literacy algorithm, it's grounded on Bayes theorem. for employed for resolving bracket issues [19]. Naive Bayes Classifier is one of the easy also veritably effective Bracket algorithms which help at constructing the quick ML ways that could produce rapid-fire vaticinators. This classifier is a probabilistic-classifier, it means it forecasts on the establishment of possibility of the thing. This algorithm is consists of 2 terms,

(i) Naive and

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 05 July 2023 **Revised:** 25 August 2023 **Accepted:** 15 September 2023*

(ii) Bayes,

which are be explained Naive then it's known as Naïve, why because it presumes the prevalence of certain attributes that's free of the prevalence.

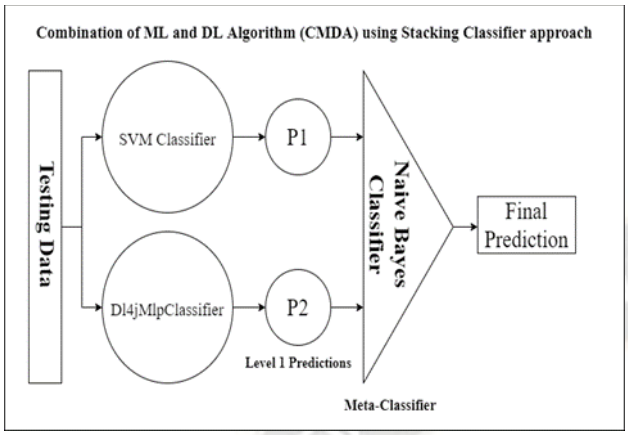

Figure 6: CMDA algorithm using stacking classifier approach

Bayes It's relies on the principle of the Baye ' Theorem. This theorem is called as Bayes ' law or Bayes ' Rule, which can be employed

$$P(A|B) = \frac{(P(B|A) * P(A))}{P(B)} \quad (2)$$

Where,

P (A|B) known as the Posterior probability: Probability of hypothesis A is calculated event B.

P (B|A) is the probability of like hood: This calculate the Probability of the evidence and it shows the particular probability of a hypothesis given is true.

P (A) is Probability of Prior: before observing the evidence the Probability of hypothesis is given.

P (B) is Probability of Marginal: It shows the Probability of the Evidence.

Risk Factor:

After heart complaint data prognosticated and uprooted, the CMD[illegible] high threat) maximum value. Also, each other is get converted between 0 and 1.

Eq. (3) describe Min-Max Normalization.

$$Ynorm = \frac{y - \min(y)}{\max(y) - \min(y)} \quad (3)$$

For case, if minimal value of a point is min(y) = 20, and the maximum value (y) = 40, also y = 30 then it would be altered to about Ynorm = 0.5. After min- maximum normalization, the CMDA algorithm computes the normal for each case data. If the normal is lower than or equal to0.3, it concludes the case has low threat. If the normal is lesser than0.3 and lower than or equal to0.6, it concludes the case has medium threat. If the normal is lesser than0.6, it concludes the case has a high threat.

5. Experimental Results & Discussions:

5.1 Performance Measurement:

5.1.1 Accuracy:

The accuracy is the more effective performance measure method taken to evaluate performance study. Accuracy is denoted in percentage of each properly forecasted data which is showed in equation 4.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

Here True Positive showed as TP, True Negative as TN, False Positive as FP and False Negative as FN.

Table2: Comparison for heart disease dataset for accuracy

Algorithm	Accuracy
SVM	0.83
Dl4jMlpClassifier	0.87
CMDA	0.96

Figure 7 describes the comparison of heart disease accuracy based on the dataset. The result is then compared with other classification algorithms. The proposed CMDA shows the high results when compared with other methods.

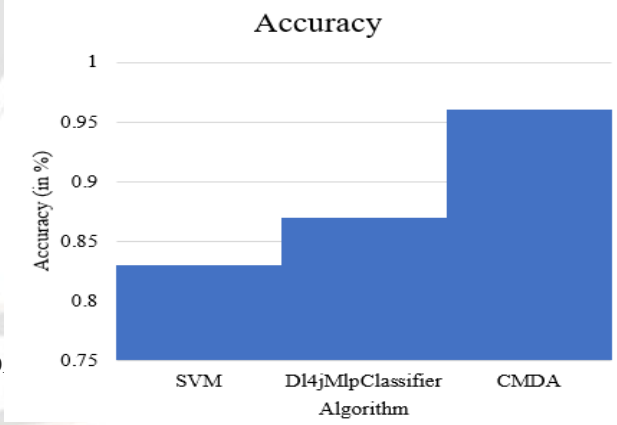

Figure 7: Accuracy Comparison for heart disease dataset

5.1.2 Precision:

The Precision is known as the measure of exactness or quality or positive predictive value.

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

Here True Positive shows TP and False Positive shows FP.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *05 July 2023* ***Revised:*** *25 August 2023* ***Accepted:*** *15 September 2023*

Table 3: Precision comparison for heart disease dataset

Algorithm	Precision
SVM	0.79
Dl4jMlpClassifier	0.94
CMDA	0.95

Figure 8 shows precision by comparison of heart disease dataset. The proposed algorithm is compared with few other classification methods. The proposed CMDA algorithm shows the high results of precision.

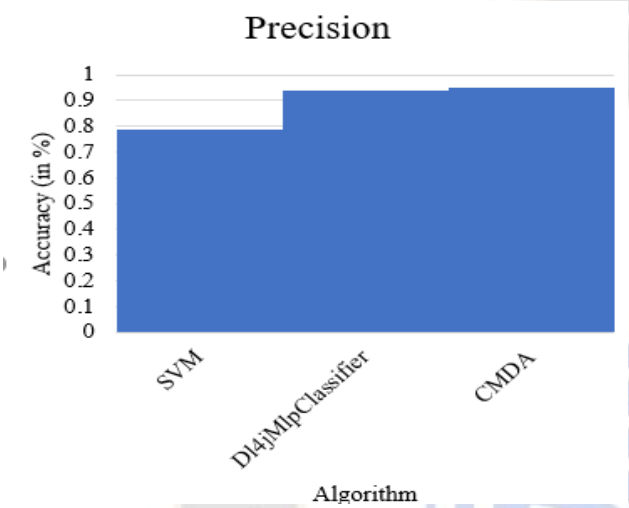

Figure 8: Comparison of Precision for heart disease dataset

5.1.3 Recall:

The recall is called as sensitivity or true positive rate (TPR) or probability detection.

$$Recall = \frac{\text{TP}}{\text{TP} + \text{FN}} \qquad (6)$$

Here True Positive means TP False Negative means FN.

Table 4: Recall comparison for heart disease dataset

Algorithm	Recall
SVM	0.78
Dl4jMlpClassifier	0.81
CMDA	0.86

Figure 9 describes the comparison of heart disease recall based on the dataset. The result is then compared with other classification algorithms. The proposed CMDA shows high results when compared with other algorithms.

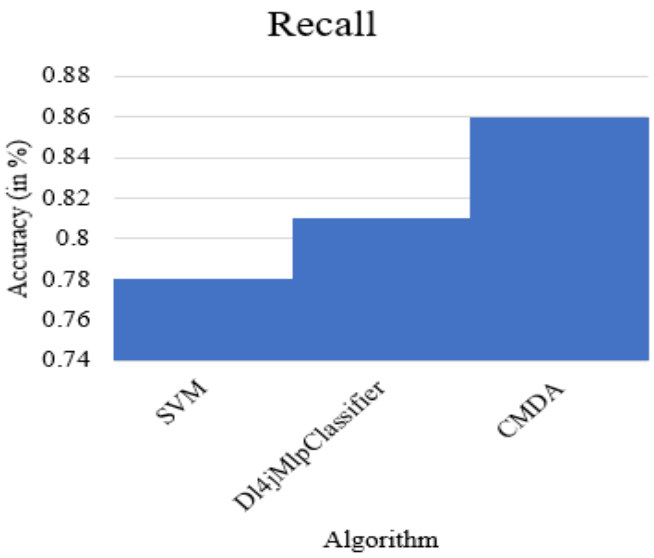

Figure 9: Comparison of Recall for heart disease dataset

5.1.4 F-Measure:

A F-Measure merges precision and recall. It is the mean of precision, and the recall both together is called as F-Measure.

$$FMeasure = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \qquad (7)$$

Table 5: F-Measure comparison for heart disease dataset

Algorithm	F-Measure
SVM	0.78
Dl4jMlpClassifier	0.87
CMDA	0.90

Figure 10 describes the comparison of heart disease F-Measure based on the dataset. The result is then compared with other classification algorithms. The proposed CMDA gives high results than other algorithms.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 05 July 2023 **Revised:** 25 August 2023 **Accepted:** 15 September 2023*

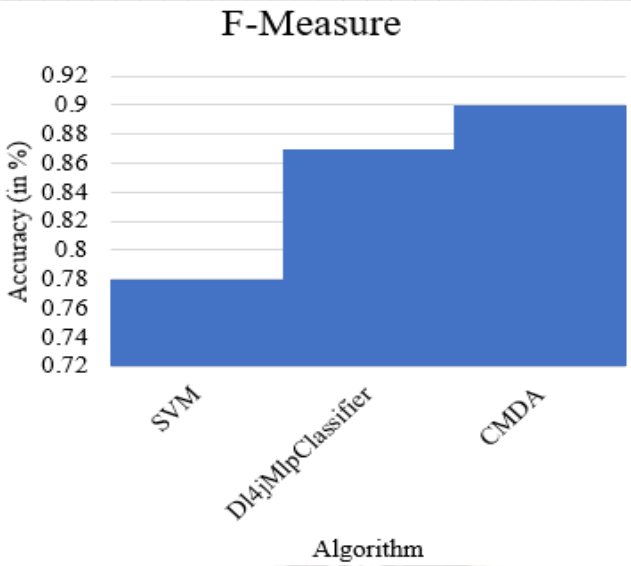

Figure 10: Comparison of F-Measure for heart disease dataset

5. Conclusion & Future Work:

This paper is proposed a work by combination of ML and DL algorithm (CMDA) for diabetes and heart disease classification and prediction. Support Vector Machine algorithm used for ML and Dl4jMlpClassifier algorithm used for DL in the CMDA algorithm. In the CMDA algorithm, the stacking classifier approach is used to combine the two algorithms. This stacking classifier is a technique where the output of the classifiers is sent as the input to the meta-classifier for the final prediction work—the Naive Bayes classification algorithm are used as the meta-classifier in the CMDA. At last, the CMDA algorithm predicts risk factor using the Min - Max normalization technique. To evaluate the performance of the CMDA algorithm, the four-essential metrics, namely precision, accuracy, recall and F-Measure. The result was concluded and it is compared with few other classification algorithms, the proposed CMDA provided with high results compared to other. In Future, the CMDA algorithm could apply to numerous research domains like various healthcare, analytics of finance, data analysis of supply chain, etc.

REFERENCES

1. Nowson, C. A., Appleton, J., & Grieger, J. A. (2018). "The impact of dietary factors on indices of chronic disease in older people". The journal of nutrition, health & aging, 22(2), 282-296.
2. Jain, D, & Singh, V. (2018). "Feature selection and classification systems for chronic disease prediction". Egyptian Informatics Journal, 19(3), 179-189.
3. Battineni, G., Sagaro, G. G., Chinatalapudi, N., & Amenta, F. (2020). "Applications of machine learning predictive models in chronic disease diagnosis". Journal of personalised medicine, 10(2), 21.
4. Fritz, B. A., Cui, Z., Zhang, M., He, Y., Chen, Y., Kronzer, A., ... & Avidan, M. S. (2019). "Deep-learning model for predicting 30-day postoperative mortality". British journal of anaesthesia, 123(5), 688-695.
5. Sujatha, R., Chatterjee, J. M., Jhanjhi, N. Z., & Brohi, S. N. (2021). "Performance of deep learning vs machine learning in plant leaf disease detection". Microprocessors and Microsystems, 80, 103615.
6. Raghavendra, S., & Santosh, K. J. (2020). "Performance evaluation of random forest with feature selection methods in prediction of diabetes". International Journal of Electrical and Computer Engineering, 10(1), 353.
7. Kadhm, M. S., Ghindawi, I. W., & Mhawi, D. E. (2018). "An accurate diabetes prediction system based on K-means clustering and proposed classification approach". International Journal of Applied Engineering Research, 13(6), 4038-4041.
8. Ayon, S. I., & Islam, M. (2019). "Diabetes Prediction: A Deep Learning Approach". International Journal of Information Engineering & Electronic Business, 11(2).
9. Mohan, S., Thirumalai, C., & Srivastava, G. (2019). "Effective heart disease prediction using hybrid machine learning techniques". IEEE Access, 7, 81542-81554.
10. Mathan, K., Kumar, P. M., Panchatcharam, P., Manogaran, G., & Varadharajan, R. (2018). "A novel Gini index decision tree data mining method with neural network classifiers to predict heart disease". Design automation for embedded systems, 22(3), 225-242.
11. Chintan M. Bhatt, Parth Patel, Tarang Ghetia and Pier Luigi Mazzeo. "Effective Heart Disease prediction using Machine Learning Techniques". Algorithms 2023.
12. Apurb Rajdhan, Apurb Rajdhan, Avi Agarwal, Dundigalla Ravi, Dr. Poonam Ghuli. "Heart Disease Prediction using Machine Learning". International Journal of Engineering Research & Technology (IJERT) 2020.
13. Harshit Jindal, Sarthak Agrawal, Rishabh Khera, Rachna Jain Preeti Nagrath. "Heart Disease Prediction using Machine Learning algorithms". Materials Science and Engineering.
14. Arsalan Khan, Moiz Qureshi, Muhammad Daniyal,3 and Kassim Tawiah. "A Novel Study on Machine Learning Algorithm-Based Cardiovascular Disease Prediction". Hindawi Health & Social Care in the Community Volume 2023, Article ID 1406060, 10 pages
15. Tintu George and A. Hema, "Data Preprocessing and Tier Feature Selection for
Diabetes and Heart Disease Data", Journal of Xidian University. Pages -189 to 191
16. M. S. Amin, Y. K. Chiam, and K. D. Varathan, "Identification of significant features and data mining techniques in predicting heart disease," Telematics and Informatics, vol. 36, pp. 82–93, 2019.
17. Sivasankari, S. S., Surendiran, J., Yuvaraj N., Ramkumar, M., Ravi, C. N., and Vidhya, R. G. (2022). "Classification of diabetes using multilayer perceptron." in 2022 International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE). IEEE, 1–5.

18. Mohan, S, Thirumalai, C, and Srivastava, G. "Effective heart B disease prediction using hybrid machine learning techniques". *IEEE Access*. (2019) 7:81542–54. doi: 10.1109/ACCESS.2019.2923707.
19. Hassan, C.A.U, Iqbal J, Irfan R, Hussain S, Algarni A.D, Bukhari S.S.H, Alturki N, Ullah, S.S. "Effectively Predicting the Presence of Coronary Heart Disease Using Machine Learning Classifier". Sensor 2022,22,7227
20. Bhunia, P.K, Debnath A, Mondal P D E M Ganguly, K Rakshit P, "Heart Disease Prediction using Machine Learning." International Journal of Engineering Research Technology. 2021
21. Harshit Jindal, Sarthak Agarwal, Rishabh Khera, Rachna Jain and Preeti Nagrath "Heart Disease Prediction using Machine Learning algorithms".Materials Science and Engineering 2021.
22. Md Juliker Nayeem, Sohel Rana, Md. Rabiul Islam, "Prediction of Heart Disease using Machine Learning algorithms". IEEE 2022.
23. Md Juliker Nayeem, Sohel Rana, Md. Rabiul Islam, "Prediction of Heart Disease using Machine Learning algorithms". IEEE 2022.
24. Yilmaz R,Yagin FH,"Early detection of coronary heart disease based on machine learning methods", International Journal of Medical Journal 2022, doi: 10.37990/medr.1011924.