International Journal of Engineering & Technology, 7 (2.33) (2018) 1040-1045

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

An Efficient Heart Disease Prediction System Using Modified Firefly Algorithm Based Radial Basis Function with Support Vector Machine

M. Thiyagaraj[1]*, Dr.G.Suseendran[2]

[1]*Ph.D., Research Scholar,Department of Information and Technology, VelsInstitute of Science, Technology & Advanced Studies (VISTAS), Chennai, India.*

[2]*Assistant Professor, Department of Information and Technology, School of Computing Sciences, VelsInstitute of Science, Technology & Advanced Studies (VISTAS), Chennai, India, Chennai, India.*

E-mail:suseendar_1234@yahoo.co.in

** Corresponding author E-mail: mthyagarajphd@gmail.com*

Abstract

Nowadays Heart Disease is one among the main roots of death in and around countries. Accurately predicting the heart disease is difficult for doctors. Thus, it is obligatory to apply computerized technologies to support doctors for diagnose faster with greater accuracy. An existing work introduced a heart disease diagnosis system which is dependent upon Interval Type-2 Fuzzy Logic System (IT2FLS). However, the training time of IT2FLS as well as genetic hybrid algorithms is quite gentle. And also it does not achieve high recognition accuracy. To solve this problem the proposed system introduced a modified firefly algorithm and Radial Basis Function based Support Vector Machine (MFA and RBF-SVM). An input dataset encompasses 3 kinds of attributes such as Input, Key and Prediction attributes. After the normalization, an attribute reduction and feature extraction are performed by using FA and Principal Component Analysis (PCA) respectively. Finally RBF-SVM is classified a features as normal or heart diseases.

Keywords:*Firefly algorithm; heart disease; normalization; support vector machine and attribute reduction.*

1. Introduction

Data mining is method of demanding understood, beforehand obscure and future valuable data from accessible information. In short it is procedure of breaking down information from various perspective and gathering the learning from it [1]. To extricate obscure examples and connections from colossal databases, information mining blends the factual investigation, machine learning and database innovation. The gain information can be helpful for various application in particular human services industry. These days, wellbeing infections are increasinggradually because of means of life, genetic [2]. Mainly, coronary illness has become more distinctive in nowadays that is to say the life of individuals is in risk. The heart is a critical organ of human body. On the off chance that the blood course to the body is deficient, the organs of the body that is mind and heart quit working and passing happens in couple of minutes. The hazard factors related are distinguished as family history, age, hypertension, diabetes, elevated cholesterol, tobacco, smoking, liquor consumption, stoutness, physical dormancy, chest torment compose and less than stellar eating routine [3].

With a specific end goal to diminish the danger of Heart Disease, forecast should be completed. Finding of Heart Disease is normally in consideration of side effects, physical examinations and indications of patient body. Regularly, specialists are anticipating coronary illness by information and experience. Finding and foreseeing illnesses is a troublesome assignment in medicinal condition. By and large, determination is by and large in light of dad tient's present test reports and the specialist's understanding. In this manner the finding is a perplexing errand that requires high expertise and much experience [4]. In the event that the coronary illness is distinguished at beginning time as well as the patient is given suitable and satisfactory treatment at appropriate phaseafterwards the life of huge amount of individuals could be spared as well as furthermore the price of the treatment could be lessened altogether. As a result there is a requirement to build up a forecast framework to identify the nearness or nonappearance of coronary illness in the patient containinggreater precision.

Coronary illness expectation framework plans to misuse the different information mining systems on medicinal informational collection to aid the forecast of the coronary illness. Datasets of coronary illness patients can be gathered from different Universities like UCI, Cleveland, and so on for our knowledge framework. The traits like age, sex, chest torment, resting circulatory strain, cholesterol mg/dl, glucose, most extreme heart rate, and so forth. With the intension of expecting the risk of coronary illness,diverse information mining techniques, e.g., Naïve Bayes, KNN calculation, Decision tree, Neural Network are utilized [5]. The KNN calculation employs the K client described an incentive to discover the estimations of the variables of coronary illness. The Naïve Bayes strategy is applied to predict the coronary illness viaprobability. Choice tree calculation is used to provide the characterized answer to the coronary illness. The Neural Network provides the limited blunder of the expectation of coronary illness. In this previously mentioned methods the patient records are grouped and

anticipated constantly [6]. The patient action is observed consistently, if there is any progressions happen, at that point the hazard level of malady is educated to the patient and specialist. The specialists can anticipate heart maladies at a prior stage as a result of machine learning calculations and with the assistance of PC innovation [7] .

The mainobjective to enhance the malady forecast precision the past framework presented a hereditary calculation based fluffy rationale approach [8]. The framework utilizes Cleveland coronary illness database. There were 13 qualities associated with the expectation of coronary illness. The hereditary calculation is utilized to decide the traits that contribute more towards the analysis. By hereditary hunt, 13 credits are decreased to 6 properties. These properties are used for characterization.

2. Literature Review

P.K. Anooj [9] exhibited a weighted fluffy run based Clinical Decision Support System (CDSS) for PC helped determination of the coronary illness. The presented CDSS for chance forecast of the heart patients encompasses two stages, for example, age of weighted fluffy principles and creating of a fluffy govern based choice emotionally supportive network. Here, information preprocessing was connected on the coronary illness informational collection for evacuating the uproarious data and to discover missing qualities. After that utilizing the incessant trait classes, the deviation extend and important qualities were figured in this technique. As indicated by the deviation go, the traits were chosen when any variationsubsists or not and furthermore the deviance extend was utilized to develop the choice tenets, who were filtered in the learning database to discover its recurrence. According to its recurrence the weight age was ascertained for each choice control got and by the assistance of fluffy participation work, the weighted fluffy tenets were acquired. The programmed technique to produce the fluffy guidelines was preference of the composed framework as well as the weighted method presented in the planned work was extra favorable position for successful learning of the fluffy framework. So as to fabricate the CDSS utilizing Mamdani fluffy surmising frame work,t hese weighted fluffy principles were utilized.

Shantakumar B. Patil [10] connected productive technique for the extraction of critical examples from the coronary illness distribution centers for heart assault forecast. In this right off the bat the information distribution center is pre-prepared considering the finalobjective to make it appropriate for the mining procedure as well as besides the K-mean grouping calculation was connected for bunching the coronary illness stockroom. Henceforth the repetitive examples pertinent to coronary illness are mined with the MAFIA calculation from the information separated. Furthermore, based on registered noteworthy weightage, the examples important to heart assault anticipation are selected. For the successful expectation of Heart Attack,the neural system is prepared with the chosen critical examples[18].

Jayshril et al [11] exhibit an expectation framework for coronary illness utilizing Learning vector Quantization (LVQ) neural system algo-rithm. The neural system in this framework acknowledges 13 clinical highlights as info as well asforesees that there is a nearness or nonattendance of coronary illness in the patient, alongside various execution measures. Dataset is gathered from Cleveland database. The primary motivation behind utilizing LVQ is that it creates models that are whatsoeveron the other handhard to translate for professionals in the individual application area. LVQ arrange is a closest neighbor design classifier in view of aggressive learning. In this work, the framework comprises of two stages, in the initial step 13 clinical traits are acknowledged as info and after that the preparation of the system is finished with preparing information by LVQ calculation. It gives the nearness or nonattendance of coronary illness. To enhance the execution of the expectation framework, it is prepared with various number of neurons and furthermore the quantities of preparing ages are fluctuated.

Jabbar et al [12] displayed a sluggish information digging approach for coronary illness arrangement. Connected data driven credit measure PCA to create class affiliation rules. So asto anticipate the event of coronary illness,this class affiliation guidelines would be utilized. Dataset is gathered from UCI store. Acquainted arrangement is a run based novelmethodthat coordinates affiliation run mining and grouping. It chooses a little arrangement of great standards and utilizations these guidelines for forecast. Languid affiliated arrangement strategy actuates class affiliation rules particular to test case. The languid learning approach extends the preparation information, just on those highlights in the test information. For the most part Apriori based run age calculation creates 2k - 1 rules for the informational collection with k things. As a result it prompts greater calculation cost. So as to diminish the number of standards produced as well as to enhance exactness utilized data driven property PCA in sluggish cooperative classification.PCA is dimensionality decrease method utilized to locate another arrangement of properties that better catches the inconstancy of the information.

Raghu et.al [13] planned a Decision Support in Heart Disease Prediction System (DSHDPS) utilizing Naive Bayes. Utilizing medicinal profiles, e.g., sex, age, circulatory strain and glucose it could anticipate the likelihood of patients acquiringa coronary illness. It is im-plemented as electronic poll application. It can serve a preparation apparatus to prepare attendants and therapeutic understudies to determine dad tients to have coronary illness. Medicinal dataset was gotten from the Cleveland Heart Disease database. With the assistance of the dataset, the examples huge to the heart assault forecast are separated[19]. The records were part similarly into two datasets: preparing dataset and testing dataset. To stay away from inclination, the records for each set were chosen haphazardly. The characteristic "Analysis" is distinguished as the anticipated property with esteem "1" for patients with coronary illness and esteem "0" for patients with no coronary illness. "PatientId" is utilized as the key; the rest are input qualities. Gullible Bayes classifier gains from the "proof" by figuring the connection between's the objective (i.e., subordinate) and other (i.e., autonomous) factors. The Naive Bayes Classifier method is especially well-matched while the dimensionality of the data sources is greater.

3. Proposed Methodology

The proposed framework planned an adjusted firefly calculation and Radial Basis Function based Support Vector Machine (MFA and RBF-SVM) for heart illnesses forecast. Maladies underneath the coronary illness umbrella integrate vein ailments, e.g.,heart mood problems, coronary supply route sickness; [14] [15]. The dataset consist of 3 kinds of properties. Info, Key and Prediction traits. Usually utilized qualities, e.g., Blood weight, Gender, Age, Cholesteroland Pulse rate are taken as info properties of which age as well as sexual orientation are non-changeable characteristics. Age is persistent as well as dynamic naturally where sexual orientation is static as well as steady. The point by point outline of coronary illness conclusion framework comprises of four noteworthy states: Normalization, trait lessening, include ex-footing and arrangement[20].

3.1. Normalization Using Min–Max Normalization

Institutionalization is associated with avoid numerical inconveniences in the midst of computation process, and notwithstanding turn away incorporate qualities in grater numeric degrees from

overpowering those in more diminutive numeric scopes. In this investigation,awell-known methodology, min– max institutionalization is associated. Min– max institutionalization maps a regard v of extraordinary dataset to v' in extent of [new_ min; new_ max] by finding out eq (1),

$$v' = \frac{v \ -v_{min}}{v_{max}-v_{min}} \times [new\_max \quad - new\_min \quad] + new\_min; \qquad (1)$$

Here, new _min, new _max define the extent of progress dataset. In this examination, New_ min =0 as well asnew_max= 1 are connected. After institutionalization, changed datasets are used for property diminishing system

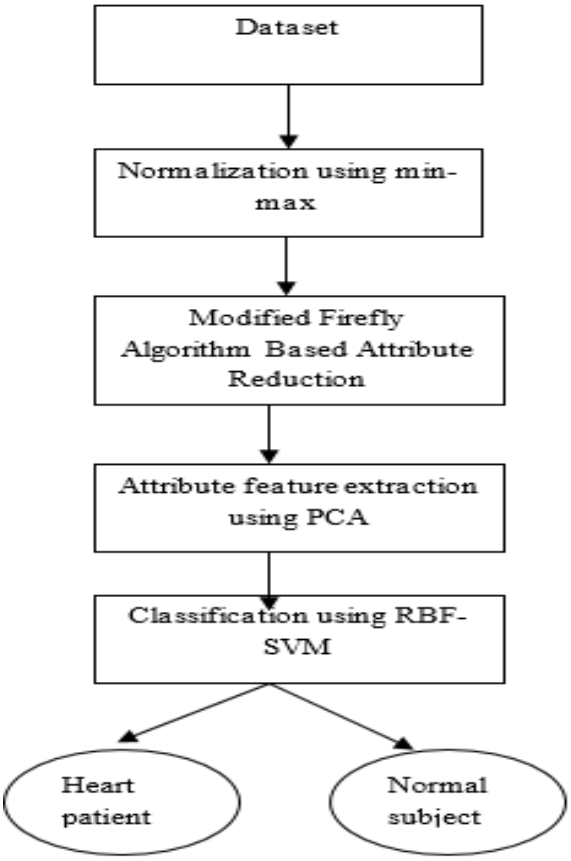

Fig. 1. Flow diagram of the proposed methodology

3.2. Modified Firefly and Rough Sets Based Attribute Reduction

There are numerous systems in light of meta-heuristic calculations and unpleasant sets werepresented to some extent writing to discover property decrease. In those techniques, the wellness capacities in light of nature of estimation of arrangement. An answer was most astounding estimation of wellness work. The estimation of nature of guess of classifi-cation in the arrangement may not equivalent to 1. Subsequently, the arrangement may not a property decrease. Likewise, those wellness capacities may not think about insignificant property lessening. To take care of this issue, this examination proposes changed firefly calculation and harsh sets based quality decrease. The brightest firefly is a firefly with current worldwide best arrangement. In the event that this brightest firefly moves arbitrarily as in the standard firefly calculation, its brilliance may diminish contingent upon the course. This prompts the decrement of the execution of the calculation in that specific emphasis. Notwithstanding, if this brightest firefly is permitted to move just if toward a path in which its splendor enhances, it won't diminish the execution of the calculation as far as worldwide best answer for the arrangements in that specific cycle. The alteration proposed in this work is as per the following. To decide the development course of the brightest firefly, haphazardly create m unit vectors, say u1, u2, . . . , um. At that point pick a course, U, among the arbitrarily created m bearings in which the splendor of the brightest firefly increments if the firefly moves toward that path..

\The primary thoughts of the firefly calculation is translating light power attributes as takes after [16]

•All fireflies are unisex and there may be a charming in any two fireflies.

•Their connecting with quality is comparing to their light power. A firefly containing cut down light power would push in the direction of the fireflies containinggreater light power. In case there isn't firefly with higher light power, the firefly would hap-hazardly move in look for space.

•By wellbeing work,the light power of a firefly is controlled.

3.2.1. Attraction Function and Movement of Fireflies

The fundamental construction of the engaging quality capacity β can be any monotonically diminishing capacity, for example, the accompanying summed up frame:

$$\beta = \beta_0 \ \times e^{-\gamma r \ \ ^2_{ij}} \qquad (2)$$

Here, r is expel amid two fireflies I and j individually. In this examination,Euclidean partition is associated. β_0 is an interest parameter, which is the drawing in quality at r = 0, c is a light ingestion coefficient.

The interest parameters as well as light osmosis coefficient are dual basic parameters. The speed of meeting and the direct of firefly computation are chosen bythe estimations of those parameters. Course of action quality has been made progress by means offine-tuning these parameters utilizing strife layout.

The Gaussian map for the attraction parameter is identified in Eq. (3)

$$\beta_0(t+1) = \left\{ \begin{matrix} 0 \quad if \quad \beta_0(t) = 0 \\ \frac{1}{\beta_0(t)} - \frac{1}{\beta_0(t)} \ otherwise \end{matrix} \right\} \qquad (3)$$

Here, β_0 (t) is known as the estimation of interest parameter at time step t. β_0 is arbitrarily created in reliably coursed in [0,1],at t= 0. In (0,1), this guide similarly makes scattered plans. Advancement of fireflies. The spots of fireflies containing cut down light power push in the direction of fireflies containinggreater light power by Eq. (4).

$$X_i(t+1) = X_i(t) + \beta X_i(t) - Xj(t) + \propto (rand - \tfrac{1}{2}) \qquad (4)$$

Here X_i (t)and X_ j(t) are spots of firefly I containingcut down lightintensity and firefly j containinggreater power at time t independently, is a sporadic parameter that chooses subjectively lead of advancement of fireflies, rand is a discretionary number generator reliably passed on in [0, 1].
In this work, every subset of trademark isconsidered as a point or position of a firefly in look for space. Wherein, every property in subset is an estimation of chase space. The perfect position is the subset of characteristics containingsmallest length as well as most vital request quality. The targets of fireflies are to movement to the finest position. There are three fundamental thoughts of use of tumult firefly figuring in quality diminishment. The primary idea is the encoding technique that is to say, to address a property di-

minish by a firefly. Second basic idea is health work. Another contemplation is parallel code of the balanced firefly estimation for property diminish issue[17].

3.2.2. Encoding Method

A possible course of action of the issue (a quality de-wrinkle) is addressed by a firefly as twofold strings of length (m =|c|). In which, each piece addresses a trademark, the esteem '1' outlines that the comparing property is chosen while the esteem '0' exhibits that the relating trait isn't chosen. For instance, assume that C={ a1, a2, . . . , a10 } and a firefly X = 1110000110, at that point a property diminishment is {a1,a2, a3, a8, a9}

3.2.3. Fitness Function

These wellness capacities may considered by rightness of characteristic lessening without insignificant diminishments . With a specific end goal to locate the best negligible characteristic decrease, it is important to apply a wellness work that not just thinks about diminishment of trait (nature of guess characterization) yet in addition thinks about insignificant property lessening. For this reason, a wellness work, is characterized in Eq. (5), is connected in this exploration.

$$\text{Fitness (X)}=\frac{m-|X|}{m}+\frac{n|R|\gamma_X(D)}{m\Gamma} \quad (5)$$

Here m =|C|, |U|; γ_x (D) is known as the nature of arrangement. R is called a reduct of condition trait C and R is registered by means of a regulate speedy decrease calculation.

This formula suggests that the course of action quality, γ_x (D, and length of attributes subset, |X|, contain various significance for characteristic reducing task. By this wellbeing work,the tolerability of position of every firefly is assessed. The objective of firefly is to movement to location with supported wellbeing regards.

3.2.4. Binary Code of Firefly Algorithm for Attribute Reduction Problem

In property decrease issue, twofold number 1 and 0 are utilized to show that relating trait is or isn't chosen. Along these lines, harsh sets based trait lessening is a discrete issue. The first firefly calculation can manage ceaseless issues. Consequently, changes of this calculation are expected to empower it to deal with parallel issue. According to the firefly count, the Equation (4) depicts firefly I with cut down light power moving in the direction of firefly containing higher light power, the circumstance of firefly I varies as an honest to goodness number.

In quality decrease issue, double number 1 and 0 are utilized to show that relating property is or isn't chosen. Along these lines, unpleasant sets based characteristic diminishment is a discrete issue. The first firefly calculation can manage consistent issues. Along these lines, changes of this calculation are expected to empower it to deal with paired issue. According to firefly estimation, the Equation (4) demonstrates firefly I containing cut down light power moving in the direction of firefly j containinghigher light power, the circumstance of firefly I changes as an honest to goodness number. As a result this number ought to be supplanted by a twofold number.

Algorithm 1: Pseudo Code of the Modified Firefly Based Attribute Reduction Algorithm
Input

Decision table DT =(U, A =CD) where C ={1, 2,. . . . …m) a set of condition attributes;
D: decision attribute; U ={ x1, x2,. . . ; xn} is non-empty finite set of objects;
Output

S⊆C; //S is the best minimal attribute reduction
1. Generate initial population of fireflies
P = Xi(1; 2; . . . ; n) ;
2. Target work F(Xi), Xi = {x1; x2; . . . ; xm);
3. Light power Ii at Xi is dictated by F(Xi);
4. Definition the firefly calculation parameters: , γ;
5. Register a characteristic diminishment in view of directed fast reduce calculation
6. while (t Max Generations) do
7. Tune alluring parameter, β_(0) (t), in light of Eq. (3);
8. Assess light force of fireflies utilizing target work utilizing eq (5)
9. Ranking light force of fireflies;10. Find the current best solution;
11. For i = 1 to n do
12. For j = 1 to n do
13. If Ii >Ij then
14. Compute attractiveness function using eq (2)
15. Vary the position of firefly I in m-dimension using eq (4)
16. Compute value of tanh function, $f(X_i^k)$

$$f(X_i^k)=\frac{\exp\left(2X_i^k\right)-1}{\exp\left(2X_i^k\right)+1}\ \text{i=1,…..n; k=1,….m} \quad (6)$$

17. If f $F(X_i^k)$ rand≥thenX_i^k else $X_i^k = 0$end if;
18. end if
19. end for
20. end for
21. t= t + 1;
22. end while
23. Post process results;

3.3. Principal Component Analysis (PCA)

When all is said in done, the PCA system changes vectors from a - dimensional space to vectors in another, - dimensional space. It extricates diminished dimensional element subset through Principal Component Analysis (PCA), whichis an element extraction method, which produces novel highlights that are straight blend of the underlying highlights. It maps all cases of the provided dataset display in a d dimensional space to a k dimensional subspace with the end goal that k < d. The arrangement of k noveldimensions produced are known as the Principal Components (PC) and every chief part is coordinated in the direction of most extreme change barring the difference as of now represented in all its first segments. In this way, the primary segment covers the maxi-mum difference and every part that tails it covers lesser estimation of fluctuation. The Principal Components can be spoken to as the accompanying

$$PC_i = a_1X_d + a_2X_d + \cdots + a_dX_d \quad (7)$$

Here,
PC_i – Principal Component 'i
X_j – original feature 'j';
a_j – numerical coefficient for X_j.
Calculation of the principal components isprovided with the subsequent algorithm:
1. Compute the covariance matrix S from the input data.
2. Calculate the eigenvalues as well as eigenvectors of S and sort them in a downward order w.r.t the eigenvalues.
3. Produce the real transition matrix by means of considering the predefined number of components (eigenvectors).
4. Lastly, multiply the original feature space with the attained transition matrix that produces a lower- dimensional representation.

3.4 Radial Basis Function-Support Vector Machine (RBF-SVM)

The RBF based SVM contains two classes that is to say Normal Subject (NS) and Heart Patient (HP).Support vector machines (SVMs) are the accumulation of regulated learning procedures utilized for dissecting the information as well as perceiving the examples. The standard SVM is a paired classifier anticipating one among the two conceivable classes as a result. It is a straight classifieras well asa non-probabilistic. The choice capacity for directly detachable issues is spoken to by:

$(\omega . x_i^{"} + b) \leq -1 \; if \;\; c_i$ =-1 and

$(\omega . x_i^{"} + b) \geq \; 1 \; if \;\; c_i$ =1 (8)

thatcould be merged covering both cases that is to say, $\omega . x_i^{"} + b) \geq 1$

Here, x" is the information, y speak to the class, b as edgeandw as weight. The issue of coronary illness analysis containing highlights anticipated onto the bearings of eigenvectors speaking to greatercovariances is a non-direct grouping issue. The SVM choice capacity for the non-straight order is in this manner:

$$f(x^{""}) = \; \omega . \varphi(x^{""}) + b \tag{9}$$

The high dimensional information confront merging issue. To understand the meeting issues SVM bits are very useful. win the bit based SVM is in this manner:

$$\omega = \; \sum_{i=1}^{m} \alpha_i \, \varphi(x_i^{"}) \tag{10}$$

Joining the piece based portrayals, choice capacity for non-straight separation is spoken to as:

$$f(x^{"}) = \sum_{i=1}^{m} \alpha_i \, \omega . \varphi(x^{""}) . \varphi(x^{""}) \; + b \tag{11}$$

Here $\varphi(x^{""}) . \varphi(x^{""})$ is called kernel functionby $KF(x^{""}, x_i^{"})$. The decision function is in this manner:

$$f(x^{"}) = \sum_{i=1}^{m} \alpha_i \, KF(x^{""}, x_i^{"}) \; + b \tag{12}$$

3.3.1. Radial Basis Function (RBF):

RBF based piece named as Gaussian part maps the lower dimensional segment space into an unfathomable dimensional space. As soon as straightforwardly no-separable features are mapped onto higher dimensional space they consistently end up being straightly discernable. The RBF divide is portrayed a

$$KF \; (x^{""}, \; x_i^{"}) = \quad e^{(-\gamma \left| x^{""}, x_i^{"} \right| \; ^2} : \forall \tag{13}$$

Theγ is in charge of fixing the width of the Gaussian ringer molded bend. Littler the estimation of the γwidensthe bend and the other way around. The RBF piece when fused in SVM a ultimate conclusion work progresses toward becoming

$$f(x^{""}) = \sum_{i=1}^{m} \alpha_i \, e^{(-\gamma \left| x^{""}, x_i^{"} \right| \; ^2} + b \tag{14}$$

The RBF based SVM contains two classifications that is to say-Normal Subject (NS) and Heart Patient (HP)

4. Performance Evaluation

Nitty gritty accommodation rules can be found on the diary site pages. All creators are in charge of understanding these rules previously presenting their original copy.

Productive Heart Disease Prediction is the most critical technique for distinguish heart ailments. Here proposed changed FA and RBF-SVM approach and existing (IT2FLS)are thought about as far as Sensitivity, Specificity and Accuracy. The metric qualities are discovered in light of True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN).

False Positive Rate (FPR)

The level of situations where a picture was fragmented to tumor part, however in actuality it didn't.

$$FPR = \frac{FP}{FP + TN} \tag{15}$$

False Negative Rate (FNR)

The ratio of cases in which an image was segmented to non-tumor part.

$$FNR = \frac{FN}{FN + TP} \tag{16}$$

4.1. Accuracy

The weighted ratio of tumor parts in images is appropriately segmented by the measurement accuracy. It is denotedin this manner,

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \times 100 \tag{17}$$

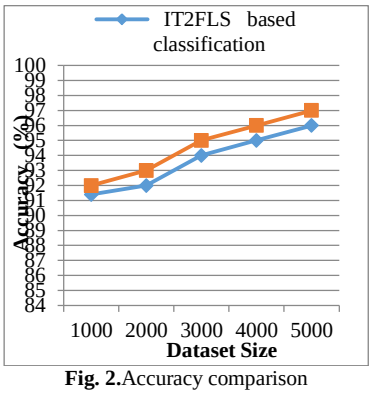

Fig. 2. Accuracy comparison

Fig 2 illustrates that the comparison of proposed MFA and RBF-SVM based classification approach and existing IT2FLS based classification approach in terms of accuracy. The size of dataset is taken as X axis as well as in y axis accuracy is taken. With the intention ofattaininggreater accuracy the proposed system used MFA is used for attribute reduction. It concludes that the MFA and based RBF-SVM classification approaches shown the high accuracy results for all size of dataset.

4.2. Sensitivity

The degree of the affectability is the extent of real positives that are appropriately perceived. It identifies with the limit of test to perceive positive outcomes.

$$\text{Sensitivity} = \frac{number \; of \; true \; positives}{number \; of \; true \; positives + number \; of \; false \; negatives} \times 100 \tag{18}$$

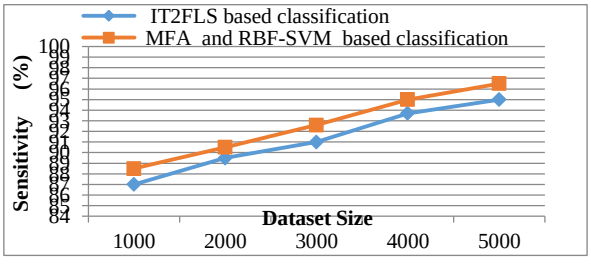

Fig. 3. Sensitivity comparison

Fig 3 illustrates that the comparison of proposed MFA and RBF-SVM based classification approach and existingIT2FLS based classification approach in terms of sensitivity. The size of dataset is taken as X axis and in y axis sensitivity is taken. In this proposed work, to improve the overall system performance min-max algorithm is used for normalization purpose. And also an efficient classification is done by using RBF-SVM. It improves the true positive rate. For all size of dataset the proposed MFA and based RBF-SVM classification approach has shown the high sensitivity results compared with the existing system.

4.3.Specificity

It isthe degree of the specificity is the ratio of negatives that are appropriately recognized. It associates with the capacity of test to identify negative outcomes.

$$\text{Specificity} = \frac{\text{Number of true negatives}}{\text{Number of true negtaives+number of false positives}} \times 100 \quad (19)$$

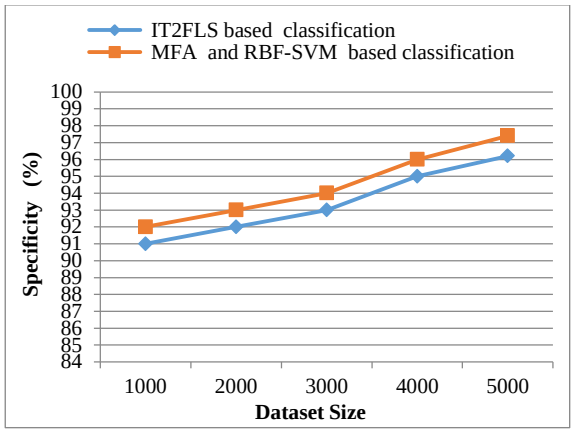

Fig. 3. Specificity comparison

Fig. 4 illustrates that the comparison of proposed MFA and RBF-SVM based classification approach and existing IT2FLS based classification approach in terms of specificity. The size of dataset is taken as x axis and in y axis specificity is taken. For all size of dataset the proposed MFA and based RBF-SVM classification approach has shown the higher specificity results compared with the existing system.

5. Conclusion

This proposed framework planned a changed firefly calculation and Radial Basis Function based Support Vector Machine (MFA and RBF-SVM) way to deal with foresee coronary illness astutely and proficiently, and to supplant manual endeavors. The given dataset comprises of three kinds of qualities to be specific, Input, Key and Prediction characteristics. In this work, standardization is performed by utilizing min– max standardization approach. At that point plan an effective heuristic approach for characteristic decrease named as MFA to oversee gigantic measures of highlights and expansive quantities of records. After that the highlights are removed by utilizing PCA. At last RBF-SVM is classi-fied a highlights as ordinary or heart sicknesses. The trial comes about demonstrate that the proposed framework accomplishes elite as far as precision, affectability and specificity. In future, register another centrality measure in view of relative perceptibility degree, and in this manner defeat the inadequacy of trait reliance. And furthermore the ongoing element extraction calculations can be proposed to enhance the general execution.

References

[1] M. Anbarasi, E. Anupriya, N.CH.S.N.Iyengar, "Enhanced Prediction of Heart Disease with Feature Subset Selection using Genetic Algorithm", International Journal of Engineering Science and Technology, Vol. 2(10), 2010, pp. 5370- 5376.

[2] Ephzibah, E. P. (2011). A hybrid genetic-fuzzy expert system for effective heart disease diagnosis. Advances in Computing and Information Technology, 115-121.

[3] Carlos Ordonez, Edward Omincenski and Levien de Braal, "Mining Constraint Association Rules to Predict Heart Disease",IEEE International Conference on Data Mining, IEEE Computer Society, ISBN-0-7695-1119- 8, pp: 433-440,2001

[4] Nilakshi P. Waghulde, Nilima P. Patil, "Genetic Neural Approach for Heart Disease Prediction", International Journal of Advanced Computer Research (ISSN (print): 2249-7277, Vol 4 Number-3 Issue-Sept 2014

[5] M. Gudadhe, K. Wankhade and S. Dongre, "Decision support system for heart disease based on support vector machine and artificial neural network", IEEE International conference on computer and communication technology, pp. 741-745, November 2010.

[6] Chaitrali S. Dangare, Sulabha S. Apte, "Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques"; International Journal of Computer Applications (0975 – 888) Volume 47– No.10, June 2012.

[7] D. Giri, U. R. Acharya, R. J. Martis, S. V. Sree, T.-C. Lim, T. Ahamed, and J. S. Suri, "Automated diagnosis of coronary artery disease affected patients using lda, pca, ica and discrete wavelet transform," Knowledge-Based Systems, vol. 37, pp. 274–282, 2013.

[8] E.P.Ephzibah, Dr. V. Sundarapandian, "Framing Fuzzy Rules using Support Sets for Effective Heart Disease Diagnosis"; International Journal of Fuzzy Logic Systems (IJFLS) Vol.2, No.1, February 2012.

[9] P.K. Anooj, "Clinical decision support system: Risk level prediction of heart disease using weighted fuzzy rules", Journal of Computer and Information Sciences, Vol.24, PP. 27–40, 2012.

[10] Shantakumar B. Patil, "Extraction of Significant Patterns from Heart Disease Warehouses for Heart Attack prediction", International Journal of Computer Science and Network Security, Vol.9, No.2, February 2009.

[11] Jayshril S. Sonawane "Prediction of Heart Disease Using Learning Vector Quantization Algorithm " IEEE publication 2014

[12] M.AkhilJabbar,Dr B.L Deekshatulu,DrPriti Chandra "Heart Disease Prediction using Lazy Associative Classification " IEEE publication 2013

[13] Dr. D. Raghu. T. Srikanth, Ch. Raja Jacub, "Probability based Heart Disease Prediction using Data Mining Techniques" IJCST Vol. 2, Issue 4, Oct - Dec. 2011, ISSN: 0976-8491 (Online) | ISSN: 2229-4333(Print)

[14] Driscoll, A., Barnes, E. H., Blankenberg, S., Colquhoun, D. M., Hunt, D., Nestel, P. J., ...& Tonkin, A. (2017). Predictors of incident heart failure in patients after an acute coronary syndrome: The LIPID heart failure risk-prediction model. International Journal of Cardiology.

[15] Saxena, K., & Sharma, R. (2016). Efficient Heart Disease Prediction System. Procedia Computer Science, 85, 962-969.

[16] Su, H., Cai, Y., & Du, Q. (2017). Firefly-algorithm-inspired framework with band selection and extreme learning machine for hyperspectral image classification. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 10(1), 309-320.

[17] D.Senthil,G.Suseendran, "Efficient Time Series Data Classification using Sliding Window Technique Based Improved Association Rule Mining with Enhanced Support Vector Machine" submitted to International Journal of Engineering & Technology, Vol. 7 (2.33) ,June 2018) pp.218-223. DOI: 10.14419/ijet.v7i2.33.13890.

[18] M.Thiyagaraj , G.Suseendran "Survey on Heart Disease Prediction System Based on Data Mining Techniques" Indian Journal of Innovations and Developments Vol 6(1), pp.1.-9, January, 2017

[19] K.Rohini, G.Suseendran, "Aggregated K Means Clustering and Decision Tree Algorithm for Spirometry Data", Indian Journal of Science and Technology, Volume 9, Issue 44, November 2016.

[20] D.Senthil, G.Suseendran , "Data Mining Techniques using Time Series Analysis" Proceedings of the 11th INDIACom; INDIACom-2017; IEEE Conference ID: 40353 2017 4th International Conference on "Computing for Sustainable Global Development", 01st - 03rd March, 2017 Bharati Vidyapeeth's Institute of Computer Applications and Management (BVICAM), New Delhi (INDIA) ISSN 0973-7529; ISBN 978-93-80544-24-3 pp.2864-2872

[21] M. Thiyagaraj , G.Suseendran, "Review of Chronic Kidney Disease Based on Data Mining Proceedings of the 11th INDIACom; INDIACom-2017; IEEE Conference ID: 40353 2017 4th International Conference on "Computing for Sustainable Global Development", 01st - 03rd March, 2017 Bharati Vidyapeeth's Institute of Computer Applications and Management (BVICAM), New Delhi (INDIA) ISSN 0973-7529; ISBN 978-93-80544-24-3 pp.2873-2878