Journal Pre-proofs

Dynamic neutrosophic cognitive map with improved cuckoo search algorithm (dncm-icsa) and ensemble classifier for rheumatoid arthritis (ra) disease

B. Chithra, R. Nedunchezhian

PII: S1319-1578(20)30411-0
DOI: https://doi.org/10.1016/j.jksuci.2020.06.011
Reference: JKSUCI 807

To appear in: *Journal of King Saud University - Computer and Information Sciences*

Received Date: 12 February 2020
Revised Date: 23 May 2020
Accepted Date: 29 June 2020

Please cite this article as: Chithra, B., Nedunchezhian, R., Dynamic neutrosophic cognitive map with improved cuckoo search algorithm (dncm-icsa) and ensemble classifier for rheumatoid arthritis (ra) disease, *Journal of King Saud University - Computer and Information Sciences* (2020), doi: https://doi.org/10.1016/j.jksuci.2020.06.011

DYNAMIC NEUTROSOPHIC COGNITIVE MAP WITH IMPROVED CUCKOO SEARCH ALGORITHM (DNCM-ICSA) AND ENSEMBLE CLASSIFIER FOR RHEUMATOID ARTHRITIS (RA) DISEASE

Author 1 : B. Chithra,
HOD, Department of Computer Technology,
Shri Nehru Maha Vidhyala College of Arts and Science,Malumachampatti, Coimbatore.
E-mail:chithra220@gmail.com
Author 2 : Dr.R. Nedunchezhian, Professor,
Department of CSE & IT,
Coimbatore Institute of Technology, Coimbatore, Tamil Nadu.
E-mail:nedunchezhian@cit.edu.in, rajuchezhian@gmail.com

DYNAMIC NEUTROSOPHIC COGNITIVE MAP WITH IMPROVED CUCKOO SEARCH ALGORITHM (DNCM-ICSA) AND ENSEMBLE CLASSIFIER FOR RHEUMATOID ARTHRITIS (RA) DISEASE

ABSTRACT: Rheumatoid Arthritis (RA) falls under the group of chronic autoimmune diseases, which affects the joints and muscles, and can lead to considerable damage to the joint structure and their functionality. RA diagnosis much in early stages is quite critical in stopping the progression of the disease. In this technical work, Dynamic Neutrosophic Cognitive Map with Improved Cuckoo Search Algorithm (DNCM-ICSA) with ensemble classifier is introduced for obtaining the gene expression profile, which distinguishes between the persons affected with RA and probable subjects of control. There are four important steps involved in this work, which include data preprocessing, feature selection, prediction and classification. The initial phase of the work comprises of data preprocessing, and second phase of the work comprises of gene selection process with T-test, chi-squared test, relief-F and Minimum Redundancy Maximum Relevance (mRMR). Next, the disease prediction is carried out using the ensemble mechanism, which increases the prediction accuracy. The ensemble mechanism integrates the process of Adaptive Neuro Fuzzy Inference System (ANFIS) and Deep Neural Networks (DNNs). The ensemble mechanism of classifiers is a group of classifiers whose decisions individually are integrated generally by weighted means for the classification of new RA examples. The disease of the patients may be avoided from getting to the severe stages. At last, DNCM-ICSA algorithm is utilized for gene classification. The results of the newly introduced classifier are analysed in terms of the metrics including precision, recall, F-measure and accuracy.

KEYWORDS: Rheumatoid Arthritis (RA), Adaptive Neuro Fuzzy Inference System (ANFIS), Deep Neural Networks (DNNs), Support Vector Machine (SVM), Particle Swarm Optimization (PSO).

1. INTRODUCTION

Rheumatoid Arthritis (RA) is a long-span autoimmune ailment, which endangers the joints chiefly. It is an inflammatory rheumatic disease with a course progressively damaging the particular and extra-particular structures leading to pain, disability and death [1]. It makes the joints warm, swollen and aching. In general, the wrist and hands are affected, with the same joints commonly present on both the sides of the body. The disease may also have an effect on the other sections of the body [2]. This may lead to a lesser count of red blood cells, inflammation around the lungs, and the heart. There may be fever and fatigue. Continuous inflammation results in corruptive damage in the joint and functional disability in several patients [3][4].

Inflammation refers to a biological process corresponding to the generation of several factors, thereby indicating that variations in the gene expression is able to render symptoms at diverse phases of the inflammatory ailments [5]. With regard to biological sciences, its application involves the gene expression data analysis. Here, a common problem is to search for a set of important genes associated with a particular phenotype. Research experts make use of gene expression data for deciding if the induction/repression of a gene is useful in the diagnosis of the disease symptoms. Few conditions of diagnosis and classification for RA have been designed in the last few decades [6] to help in the gene expression data. However, the implementation of the criteria for a diagnosis carried out in a timely manner is still under debate [7-8].

Recently, data mining employing samples has gained much popularity and is anticipated to help in increasing the accuracy involved with the diagnosis and quality of treatment employing data mining approaches. Information acquired (with the help of data mining) from the clinical data render the physicians an extra source of knowledge to make decisions in their practical work, planning of treatment, risk analysis and other kinds of predictions. The diagnosis on diseases is one among the applications where data mining has found tremendous success. The objective of the prediction model is the Design of a model, which can find the predicted class [9].

The objective of classification, supervised learning approach involves the prediction of the target class for every case in the data in which the classes are predetermined. Currently, even though different classification algorithms select the best model for a particular data set since these classical algorithms are affected by general issues, such as computation wise complexity, slipping to local minima or over-fitting to the data set utilized for training purposes [10]. One of the most popularly employed schemes forgetting over these problems involves the ensemble learning, chiefly utilized forx the performance improvement of a classifier. It involves a process in which predictions made from different classifiers are integrated with the aim of classifying new samples to improve the prediction accuracy [11].

Ensemble techniques integrating the result obtained from classifiers have been extremely helpful in the generation of predictions with better accuracy for several complicated classification tasks. Highly popular ensemble techniques help in balancing between the diverseness and accuracy achieved of the ensemble. Also, the resourcefulness of these techniques are due to their capability of having a proper consolidation of the accurate predictions and the right errors across several different base classifiers [12-13].

In this proposed article, Dynamic Neutrosophic Cognitive Map with Improved Cuckoo Search Algorithm (DNCM-ICSA) with ensemble classifier is introduced for the gene expression profile extraction. Ensemble technique integrates the process of Adaptive Neuro Fuzzy Inference System (ANFIS) and Deep Neural Networks (DNNs). These ensemble mechanisms are helpful in the RAdisease prediction. At last, this RA disease prediction is carried out using ensemble mechanism, which increases the prediction accuracy.

2. LITREATURE REVIEW

Shanmugam and Preethi[12] studied about a Machine Learning based Ensemble Analytic Approach (MLEAA), comprises of two phases, which include: learning stage and prediction stage. During the learning phase, the data processing is carried out with the help of map reduce framework in Hadoop and the attributes defined are functioning towards prediction stage. The prediction stageof MLEAA approach includes three diverse algorithms, known as AdaBoost, Support Vector Machine (SVM), Artificial Neural Network (ANN) and depending on the voting system, the final predictive value is computed. Effective results are achieved here and it will be quite helpful in the prediction ofRA much in advance.

Sharonet al [14] designed an Ensemble algorithms like bagging, AdaBoost and random subspace with a base classifier including Random Forest (RF) and SVM and were trained and tested with the help of evaluation conditions like accuracy, precision, sensitivity and Area Under the Curve (AUC) employing Weka tool. The preliminary observations conclude that the base classifier SVM offers superior classification accuracy compared to the rest and with base classifier RF, Adaboost achieved a slightly better performance compared to other models used for RA dataset.

Shiezadeh et al [15] introduced a Cuckoo Search (CS)-Boost algorithm for the performance optimization of Adaboost algorithm and helpful in the RA diagnosis. The information were gathered from affected persons, who were cited to the rheumatology clinic of Shiraz University of Medical Sciences. In the next step, the preprocessing of data is conducted. The optimum RA disease predictive model uses CS-Boost in the form of classification algorithm. The comparison analysis between the models has revealed that CS-Boost bags the highest degree of accuracy compared to the rest. The results show that elbow and knee joints, gender, number of joints and ESR test result impact the maximum in the RA diagnosis.

Bardhan and Bhowmik [16] introduced an inflammation based structural feature analysis, and during the primary step of classification, arthritis damaged knee joint thermo grams, and every other kind of thermo grams (no arthritis) are are classified with an accuracy level of 91%. The advanced RA diagnosis during the subclinical phaseis of extreme help to the physicians in decreasing the impact of the disease. During the second step of classification, the RA and non-RA categorization are integrated through the extraction of texture, shape and frequency level features. Experimental tests reveal that the mix of all these features reduces the accuracy ofRA classification detection. To improve the classification rate, the accuracy based feature selection process is included.

Shanmugam and Preethi [17] designed a Neutrosophic Cognitive Maps (NCMs) with Genetic Algorithm (GA) framework. In the first step, the pre-processing of the Arthritis data set is carried out through Integer Scaling Normalization. Using the pre-processed data, Categorical Principle Component Analysis (PCA) technique extracts the specific features. The feature extraction is based on range classification. It is carried out with the help of NCMs with GA. The disease can be predicted with ease using the classified data set and it also renders the elaborate information on the kind of Arthritis. Therefore it will of much help in the advance prediction and diagnosis of Arthritis disease.

Salmeron et al [18] suggested an advanced prediction of patients with RA employing soft computing technique of Fuzzy Cognitive Maps (FCMs). In the first step, a condition set for RA diagnosis, consultation with a group of medical experts have been determined. After this, Particle Swarm Optimization (PSO) and FCMs were helpful in modeling this problem and the RA disease's severity is calculated. In the last step, an online survey has been carried out in Shohada University Hospital, Iran for assessing accuracy of the newly introduced model.

Kourilovitch et al [19] presented a RA diagnosis that depends on the criteria of classification, which consists of four parameters such as joint involvement, serology, degrees of acute phase reactants and the time duration of the signs. This classification eases out the classification of the patients suffering with early RA, but, the diagnosis needs highly skilled experts who are capable of distinguishing between the early symptoms of RA and other pathology.

Singh et al [20] designed a system for the RA diagnosis employing Fuzzy Logic Controller (FLC), which is a popular application of fuzzy set theory formulated by Zadeh. It is an extremely useful tool for tackling with ambiguity and inaccuracy. Therefore, an FLC can be used for modelling the knowledge of a physician. The FLC

performance relies on its knowledge base comprising of a data base and also a rule base. It has been found that the FLC performance is primarily based on its rule base, and the optimization of the membership function distributions that are maintained in the data base is a process of carrying out fine tuning.

Pandeyet al [21] demonstrated a fuzzy logic system for diagnosing Rheumatic Arthritis Fever (RAF). Fuzzy logic could be used for designing a Decision Support System (DSS)application, which could prove to be helpful in diagnosing arthritis pain in four diverse phases, such as: Fairly Mild, Mild, Moderate and Severe. A knowledge-base, which was designed applying the WHO constraints for the RAF diagnosis, expert guidelines acquired from Nepal and a Mat lab fuzzy tool box act as the tools utilized for developing the system.

Kim and Tagkopoulos [29] studied an important review of the machine-learning algorithms presently used in the study of clinical data, the merits and demerits of these algorithms, and how they are able to be leveraged in the area of rheumatology. Aim is toward assist clinicians and rheumatologists toward recognize better the essentials of Machine Learning (ML) and its applicable study purpose.

Regueiro et al [30] aimed toward assess the possible of the concordant occurrence of Rheumatoid Factor (RF), anticyclic citrullinated peptide(anti-CCP) and anti-carbamylated protein antibodies toward get better present RA classification between Early Arthritis (EA) patients.

Shanmugam and Preethi [31] designed a hybrid optimization approach which is depending on the grouping of the features of Iterative Dichotomiser 3 (ID3) and Particle Swarm Optimization (PSO) designed for feature selection and classification of RA. The efficiency of the proposed diagnosis strategy is confirmed via its accuracy, specificity, sensitivity, positive predictive value and negative predictive value by means of presented methods.

3. PROPOSED METHODOLOGY

The newly introduced work comprises of four important steps, which include data preprocessing, feature selection, prediction and classification. During the initial phase of the work, data preprocessing is carried out, and the second phase of the work employs gene selection with T-test, chi-squared test, relief-F and Minimum Redundancy Maximum Relevance (mRMR). In the third step, the disease prediction is carried out using the ensemble mechanism which integrates the process of the Adaptive Neuro Fuzzy Inference System (ANFIS) and Deep Neural Networks (DNNs). At last, DNCM-ICSA algorithm is helpful in the gene classification process. The results obtained of the decisions made individually are integrated using weighted majority function.

Motivation of the work

RA diagnosis much in early stages is quite critical in stopping the progression of the disease. It motivates to introduce a new prediction and classification model for the diagnosis of RA. Data mining methods have been employing samples has gained much attractiveness and is estimated to assist in increasing the accuracy concerned by means of the diagnosis.

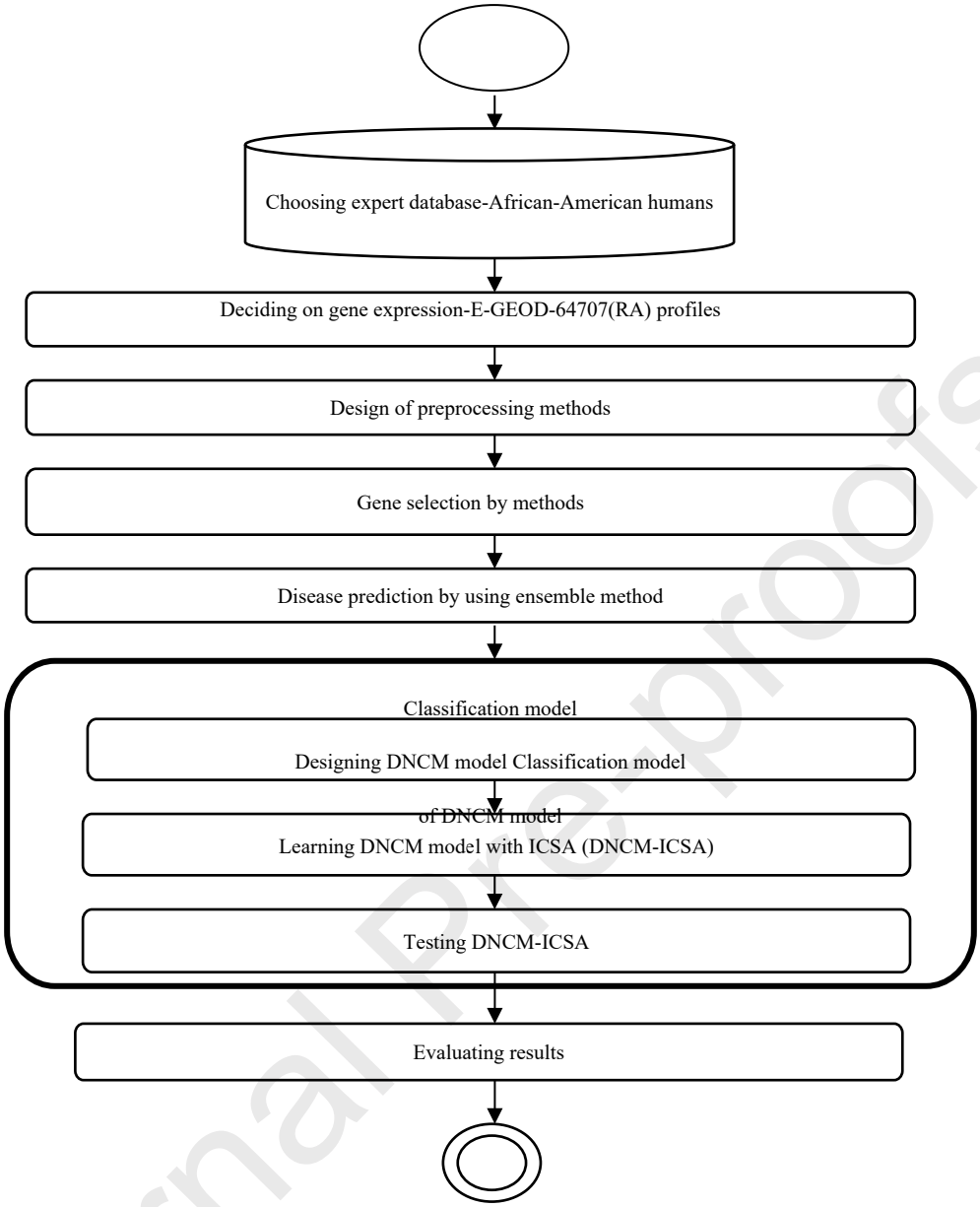

Figure 1: Architectural Diagram of Proposed DNCM-ICSA Approach

3.1. Choosing the Team of Experts

The factors behindthe radiographic harshness of Rheumatoid Arthritis (RA) in African-Americans are interpreted in a poor light. The genes whose expression in Peripheral Blood Mononuclear Cells (PBMCs) is correlated with the radio graphic criticality of RA has to be investigated. 20 control samples (obtained from persons not affected with RA) were then compared with 10 early severe, 10 early mild, 10 late mild, and 10 late severe RA samples.

3.2. Preprocessing Techniques

Normalization [22] is used in data mining system in the form of a data preprocessing component. The normalization of an attribute of a dataset is done through the scaling of its values with the objective that they are well within a specific range that is defined small, for example 0.0 to 1.0. It is primarily advantageous for learning algorithms. The mechanism of data normalization [22] consists of z-score normalization.

Z-Score Normalization

It is called as zero-mean normalization in which the values for a specific gene G get normalized based on the mean and standard deviation of gene G. A dataset (RA_D) of G gets normalized to $RA_{D'}$using the expression given below:

$$RA_{D'} = \frac{[RA_D - \text{mean}\ (G)]}{[std(G)]}\ (1)$$

Here std (G) represents the standard deviation of gene G and mean (G) indicates the average of gene G.

3.3. Gene Selection Steps

Primarily, Gene selectionapproaches are introduced for two factors:

1. Reduce the search space by removing the unnecessary variables
2. To maximize the predictive potential of the classifiers in the case of supervised learning.

Gene selection refers to the process of preprocessing the input dataset with the objective of assessing the current attributes with the objective that only the data associated with the genes are preserved and irrelevant gene data are removed. Gene selection is of use when the dimension of the dataset is big (with multiple number of attributes). The popular gene selection mechanisms (T-test, chi-squared test, relief-F and Minimum Redundancy Maximum Relevance (mRMR)) are discussed in this section.

T-test

A t-Test is fundamentally a statistic measure, which verifies whether the two mean of genes are basically diverse from one another. A bigger t-value denotes dissimilar groups and a smaller t-value represents identical groups. For every t-value, there is a respective p-value.

Chi-squared Test

This is also a well-known filter approach, which can be utilized for the gene selection process. The χ^2-Statisticvalue is computed for every gene associated with the classes individually [22]. Every numeric gene needs to be discretized before χ^2-Statistic is computed. For everygene X_i, χ2-Statistic is computed as

$$\chi^2 = \sum_{x \in X_i} \sum_{c \in C} \frac{(n_{(x \in X_i \& c \in C)} - e_{(x \in X_i \& c \in C)})^2}{e_{(x \in X_i \& c \in C)}}\ (2)$$

Where $n_{(x \in X_i \& c \in C)}$indicates the number of samples or patients belonging to class c in which $X_i$takes the value x. The required frequency $e_{(x \in X_i \& c \in C)}$is defined as

$$e_{(x \in X_i \& c \in C)} = \frac{n_{x \in X_i} * n_{c \in C}}{n}\ (3)$$

Where $n_{x \in X_i}$denotes the number of samples in which $X_i$is assigned with the value x and $n_{c \in C}$specifies the number of samples belonging to class c. n indicates the overall number of samples. The genes are chosenonce the genes are ranked in accordance with the χ^2-Statistic values.

Relief-F

Relief-F is a well-known filtering approach, which can tackle with noisy and heterogeneous datasets. The primary idea behind the Relief-F is to evaluate the quality of genes depending on their values to distinguish amongst instances, which are near to one another.

Minimum Redundancy Maximum Relevance (mRMR)

This filter mechanism chooses the genes with the most relevance and least redundancy with the target class. Maximum Relevance and Minimum Redundancy of genes are calculated on the basis of mutual information. Provided the ith gene g_i and the class label c, the mutual information of g_i and c is computed in terms of their probabilities $p(g_i)$, $p(c)$, and $p(g_i,c)$expressed as follows:

$$I(g_i,c) = \sum\sum p(g_i,c) * ln\frac{p(g_i,c)}{p(g_i)p(c)}\ (4)$$

The Maximum Relevance approach chooses the topmost m genes with the most relevance corresponding to the class labels obtained from the descending ordered set of $I(g_i,c)$.

$$\max D(S,c), D = \left(\frac{1}{|S|}\right) \sum_{g_i \in S} I(g_i,c)\ (5)$$

3.4. Ensemble Model

Ensemble is a Machine Learning approach whose techniques are meta-algorithms, which integrate various machine learning schemes into one optimal predictive model with the aim of reducing the variance, bias or increasing the predictions. This mechanism facilitates superior performance in terms of prediction in comparison with that of an individual model. This research work integrates the process of Adaptive Neuro Fuzzy Inference System (ANFIS) and Deep Neural Network (DNN).Ensemble techniques make use of two learning techniques like ANFIS and DNNs to achieve improved predictive performance compared to that achieved from any of the learning algorithms included independently. An ensemble is by itself a supervised learning algorithm, since it can undergo training and thereafter achieve the prediction capability.

3.4.1. Adaptive Neuro Fuzzy Inference System(ANFIS)

ANFIS helped in the initial membership functions computation by testing itself usingE-GEOD-64707 dataset, and then membership functions are adjusted with either a back propagation algorithm or a hybrid-learning algorithm (a mix of back propagation and the least squares technique) to reduce the error. ANFIS Architecture: The below lines define the five layers of the ANFIS system[23].

Layer 1: The primary layer includes the Fuzzification layer. Each node i present in this layer is actually a square node, expressed by:

$$O_i^l = \mu_{Ai}(x), \text{for i} = 1,2 \quad (6)$$

$$O_i^l = \mu_{Bi-2}(y), \text{for i} = 3,4 \quad (7)$$

where x and y refer to the E-GEOD-64707 dataset inputs to the node i and the fuzzy membership grade of inputs form the resultant outputs. With the aim of calculating the degree of membership of the input, each node employs the Gaussian membership function.

$$O_i^1 = \mu_{Ai}(x) = e^{\frac{-1}{2}\left(\frac{x-c_i}{\sigma_i}\right)} \quad (8)$$

where {c,σ} refers to a parameter set. C indicates the center of the membership function and σ decides the width of the membership function. These parameters are called as premise parameters.

Layer 2: The second layer includes the Rule layer. This layer takes the E-GEOD-64707 dataset input values to be the membership functions and every node multiplies E-GEOD-64707 dataset inputs and provides an output that indicates the firing strength of rule. The resultant output obtained of this layer is provided in equation

$$O_i^2 = w_i = \mu_{Ai}(X1) \times \mu_{Bi}(X2) \times \mu_{Ci}(X3) \times \mu_{Di}(X4) \times \mu_{Ei}(X5) \times \mu_{Fi}(X6) \times \mu_{Gi}(X7)\ i = 1,2,...,7$$

Layer 3: Here the i-th node is computed by dividing the i-th rules firing strength by the summation of the rule's firing strengths.

$$O_i^3 = \overline{w} = \frac{w_l}{w_1 + w'}\, i = 1,2,.....,7 \quad (9)$$

Layer 4: The fourth layer uses the nodes, which are adaptive.

$$O_i^4 = \overline{w}_i f_i = \overline{w}_i(p_i x_i + q_i x_2 + r_i x_3 + a_i x_4 + b_i x_5 + c_i x_6 + d_i x_7 + t_i) i = 1,2,...,7 \quad (10)$$

where, (w) refers to the output of layer 3, and { p_i , q_i , r_i ,...t_i}stands for the parameter set which is called as consequent parameters.

Layer 5: This layeris formed by a single fixed node. In this layer, the resultant output is given as the summation of all the inward signals, provided in the equation below

$$O_i^5 = y = \sum_i \overline{w}_i f_i = \frac{\sum_i w_i f_i}{w_i f_i}\, i = 1,2...,7 \quad (11)$$

3.4.2. Deep Neural Network (DNN)

A learning algorithm forms the model for Deep Neural Network (DNN) classification. It is helpful in modifying the weights of linear and nonlinear transfer functions present within each of the neuron units present in the network for the DNN classification model, thus attaining a required design goal generally in terms of the Minimum Mean Squared Error (MMSE) optimally. DNN classification model generally uses a regularization algorithm is L_2 regularization or weight decay and it is a direct function mechanism for regularizing the DNN classification model [24]. But, with the aim of reducing a L_2 norm, the regularization algorithm is helpful in penalizing bigger weights with the help of a hyper-parameter λ, utilized for defining the relative significance of the L_2 norm for minimizing loss faced in training dataset. DNN classification model $f_\theta(x_i)$includesgetting a weight function $\theta(w,b)$where w and b stand for the weights and bias, such that the required regularization loss is reduced [25]:

$$E(\theta,D) = \arg\min_{\theta} \left\{ \frac{1}{D} \sum_{(x_i,t_i) \in D} E(f_\theta(x_i),t_i) \right\} + \lambda \|\theta\|_p (12)$$

$x_i$indicates the inputs, and T_i are then used toλ target variable data. The hyper-parameter limits the regularization algorithm. The first function used in the regularization algorithm indicates an error function whereas the second function is called as a regularization error expressed by

$$\|\theta\|_p = \left(\sum_{j=0}^{N} |\theta_j|^p \right)^{\frac{1}{p}} (13)$$

The quality of performance of the deep neural network prediction (or diagnostic) model relies hugely on the DNN classification model while predicting the RA disease. Typically, the DNN prediction model having (L-1) hidden layers provides an output function:

$$Y = \phi_L((...\phi_3(\phi_2(\phi_1(XW_1 + B_1)W_2 + B_2)W_3 + B_3)...)W_L + B_L) (14)$$

Where the input matrix data X is sent to the layer as input; $W_n$and B_n, n = 1, 2, …, L, stands for the weight matrix and bias vectors, correspondingly, for one of the n hidden layers; the transfer n, n = 1, 2, …, L, is either a linear or nonlinear function. The final layer at Φfunction at n = L is called as an output layer, and the next other layers form the hidden layers in the DNN system and the framework. Consequently, the DNN prediction prototype can be helpful in the detection of RA disease with E-GEOD-64707 dataset during diagnostics.

3.4.3 The role of the DNNs and ANFIS :

The disease prediction is carried out using the ensemble mechanism, which increases the prediction accuracy.The ensemble disease prediction mechanism integrates the process of Adaptive Neuro Fuzzy Inference System (ANFIS) and Deep Neural Networks (DNNs). The ensemble mechanism of classifiers is a group of classifiers whose decisions individually are integrated generally by weighted means for the classification of new RA examples. Ensemble techniques make use of two learning techniques like ANFIS and DNNs to achieve improved predictive performance compared to that achieved from any of the learning algorithms included independently. An ensemble is by itself a supervised learning algorithm, since it can undergo training and thereafter achieve the prediction capability

3.4.3. Weighted Majority Voting

Hard voting can be considered to be a simpler scenario of majority voting [26]. In this, the prediction of the class label $\hat{y}$ is done through majority (plurality) voting of every classifier C_j:

$$\hat{y} = mode\{C_1(x),C_2(x),...,C_m(x)\} (15)$$

Supposing that three classifiers are integrated, which classifies a training sample as below:

- classifier 1 -> class 0
- classifier 2 -> class 0
- classifier 3 -> class 1

$$\hat{y} = mode\{0,0,1\} = 0 \qquad (16)$$

Through majority vote, the sample would be classified as "class 0."

3.5. Designing DNCM Model

Neutrosophic Cognitive Map (NCM) is called as a Neutrosophic directed graph having perceptions including rules, events and so on, in the form of nodes and causalities or in determinates to be edges. The concept on just fuzzy cognitive maps corresponds to the association/non-association amidst two nodes. It is a logic in which, every logical statement is predicted to have the proportion of truth in a subset T, and the proportion of uncertainty in a subset I, and the ratio of inaccuracy in a subset F, in which T, I, F refer to benchmarked or non-standard original subsets. SupposeC_1, C_2,.....,$C_n$indicates n nodes and the nodes represent descriptive RA concepts, which can be either features or behaviors of the system. As a result, a node C_i isrepresented by $(x_1,x_2,.....,x_k)$in which x_k's either have the value of 0or 1 or I (I refers to the indeterminate) and $x_k = 1$ indicates that the node $C_k$exists in the bigger stage of RA and $x_k = 0$ represents that the node exists in the lower stage of RA and $x_k = I$ represents that the state of nodes is in an intermediate stage of RA. Consider $C_i$and C_j indicate the two nodes (which are thenotions of RA) of the NCM. The directed edge running from $C_i$to $C_j$indicate the causality of $C_i$on $C_j$called as connections. Each edge in the NCM gets weighted with a number

present in the set {-1, 0, 1, I}. Consider W_{ij} refers to the weight of the directed edgeC_iC_j, W_{ij}∈ {-1, 0, 1, I}. W_{ij} = 0 if C_i doesn't make any consequence on C_j, W_{ij} = -1 when rise (or reduction) in $C_i$generates an increment (or reduction) inC_j, W_{ij}= 1 when increase (or reduction) in $C_i$generates an increase (or reduction) in C_j. W_{ij}= I when the association or consequence of $C_i$on C_j signifies an indeterminate.NCMs having edge weight from {-1, 0, 1, I} are known as simple NCMs. Take the Neutrosophic matrix $N(E)$isdefined as N(E) = (W_{ij}) in which W_{ij} is known as the weight of the directed edge C_iC_j, in which W_{ij}∈ {- 1, 0, 1, I}. $N(E)$is called as the Neutrosophic adjacency matrix of the NCM.Take $A = (a_1, a_2,\ldots.., a_n)$in which a_i∈ {-1, 0, 1, I}. A is called as the instantaneous state Neutrosophic vector and it signifies the on-off indeterminate state of the node at and instantofa_n.

- $a_i = 0$ if a_i is in off state (low)
- $a_i = 1$ if a_i is in on state (high)
- $a_i = I$ if a_i is in intermediate state (medium) for$i = 1,2,\ldots\ldots.,n$.

Consider C_1C_2 , C_2C_3 , C_3C_4 ,..., $C_{n-1}C_n$ be the cycle followed, if C_i gets turned on and if the causality flows through the edges in a cycle and if it generates C_i one more time, then the dynamical system goes through a cyclic pattern. This holds true for any nodec_i, for i = 1, 2,…,n. The equilibrium state for this dynamical system is known as the hidden pattern. Assume that the NCM stays with C_1 and C_n on, which is to state that the state vector continues as (1,0,…1) and this Neutrosophic state vector (1,0,…, 0,1) is called to be the fixed point. If the NCM gets settled with a Neutrosophic state vector restated as $A_1 \rightarrow A_2 \rightarrow \ldots \rightarrow A_i \rightarrow A_1$, this equilibrium is referred as the NCM's limit cycle.

3.5.1. Dynamic Neutrosophic Cognitive Map (DNCM) Model

Adhering to the objective of improving the capability of NCMs, a Dynamic Neutrosophic Cognitive Map(DNCM) model is presented, which is capable of reflecting the dynamic conducts and developing nonlinear associations in the systems. As to the adjacency matrix $N(E)$value to concept RAC_i, its relative location finds its specific weight with respect to other concepts. Consequently, W_{ij} is explained along these lines:

$$w_{ij} = \begin{cases} 0 \in domain(low) \\ 1 \in domain(large) \\ -1 \in domain(no) \\ I \in domain(intermediatestate) \end{cases} \quad (17)$$

W_{ij}takes three different values strictly. Weight learning in NCM is equivalent to the problem ofoptimization in the connection matrix, which is solved using the equation 17. NCM learning is focused on the learning of the adjacency matrix (a_i) and also on the available earlier unprocessed RA information. The learning approaches for NCMs are focused on the learning of the adjacency matrix depending on heuristic information of experts or on the available historic RA or else on both of them. Evolutionary DNCM learning approaches compute adjacency matrices from earlier information, which is a best suitable fit for the set of input state vectors. The learning goal of DNCM evolutionary learning is the generation of perfect adjacency matrices for individual NCM modeling based systems.

3.5.2. Improved Cuckoo Search Algorithm(ICSA)

Cuckoo Search (CS)

CS gets its inspiration from the obligate brood parasitism placing their data point in the cluster belonging to host birds. This minimizes the possibility of the data point being discarded and, hence, improves their reproduction capability. It is to be noted that that many host birds directly

contradict the intrusive data point. In this scenario, in case the host birds finds that the data point are not theirs, then they will either abandon them or justly eave their cluster and construct new ones, in another place. Generally, the data point hatch some time little before their host data point. At first, the instinctive action is to get the host data point out through the blind pushing of the data point out of the cluster. This act helps in improving the data point a share of food given by its host bird. The CS helps in modeling such breeding act and, therefore, can be used for different problems of optimization, and it is found that the CS performance can be enhanced with the help of Lévy Flights rather than the ordinary random walk[27]. If nature is considered, animals look out for prey either in random or semi-random fashion. Typically, the foraging path of an animal is truly a random walk as its next movement depends on both the present position/state and the transition probability to the next subsequent position. The selected direction is inherently based on a probability, and its modeling can be done analytically. Different research works have revealed that the flight behavior of several animals and insects shows the general features found in Levy flights. A Lévy flight is basically a random walk

in which the step-lengths are spread in accordance with a heavy-tailed probability distribution. Once the steps of a larger number are reached, the distance from the start of the random walk becomes a static distribution.

Cuckoo Search Implementation

Every data point in a cluster indicates a solution, and a data point indicates a solution newly found. The objective is to use the new and much superior solutions (data point) for replacing the comparably inferior solutions in the nests. Simply said, every cluster contains one data point. The algorithm can be increased in scope to more complex scenarios in which every cluster contains multiple data points indicating a group of solutions. The CS relies on three ideal rules:

- Every cuckoo generates one data point at a certain point of time, and puts it into any random cluster
- The best cluster having a higher quality of data point (solutions) will then be carried forward to the next generations
- The total number of host nests that are available is constant, and a host can find an alien data point having a probability p[a]. In this scenario, the host bird can either discard the data point or leave the cluster to construct an entirely new cluster elsewhere.

For the sake of simplicity, the final guesstimate can be approximated using a fraction a p of the n cluster getting substituted with new clusters, with newer randomized solutions. In the case of a maximization problem, the quality or fitness pertaining to a solution can just be in proportion to the objective function. Few other kinds of fitness can be expressed just like the fitness function used in genetic algorithms. In accordance with the rules stated above, the pseudo code gives the fundamental steps of the CS, as below:

Procedure of Cuckoo Search Algorithm

begin

Objective function $f(x), x = (x_1,\ldots\ldots,x_d)^T$

Create initial population of

n host clusters $x_i(i = 1,2,\ldots,n)$

While(t<MaxGeneration) or (termination criterion)

Acquire a cuckoo in random using Levy flights

assess its quality/fitness F_i

select a cluster among $n(say.j)$in random

if($F_i > F_j$)

substitute j with the new solution;

end if

A proportion(p_a) of not so good clusters

are discarded and new ones are constructed;

Retain the best solutions

(or clusters having high fitness solutions);

Sort the solutions and get the current best

end while

Post process the results and visualization

End

when new solutions $x_i(t+1)$ are generated for the i[th] data point, the Lévy flight process is carried out.

$$x_i(t+1) = x_i(t) + \alpha \oplus Levy(\lambda) \quad (18)$$

where α >0 refers to the step size, which in turn, has to be relative to the scaling in the problem of interest. The product ⊕implies multiplications done entry-wise. A Lévy flight in which the step-lengths are spread as per the probability distribution given below is defined.

$$Levy\ u = t^{-\lambda}, 1 < \lambda \leq 3 \quad (19)$$

In this step, the sequential jumps/steps of a data point generally create a random walk, which follows a power-law step-length distribution having a potentially strong tail. It is good mentioning that, practically, in the event of a data point being quite identical to the data points of the host, then this data point has the least probability of being found out, and therefore the fitness has to be associated with the difference in solutions.

Hence, it is a better idea to perform a random walk in a preferential manner using some randomly chosen step sizes.

Improved Cuckoo Search (ICS)

The parameters p_a, λ and αused in the CS assist the algorithm in finding both globally and locally extended solutions, correspondingly. The parameters $p_a$and α are quite significant parameters used in modifying the solution vectors, and can find huge application in making adjustments to the convergence rate of the algorithm. The classical CS algorithm makes use of constant value for both $p_a$andα. These values get defined during the initialization stage and cannot be modified while newer generations spring up. The primary disadvantage of this technique shows up in the number of iterations to get an optimized solution. In case the value of p_a is a small one and that ofofα is big, then the algorithm performance will be not so good and result in the number of iterations substantially increasing. In case the value of p_a is big and that of α is small, then the convergence speed becomes high, however it may be incapable of finding the best cluster. One primary variance between the ICS and CS lies in the way p_a and α are adjusted. In order to boost the performance achieved of the CS algorithm and get over the disadvantages with static values of p_a and α , the ICS algorithm makes use of variables p_a and α.During the prime initial generations, the values of p_a and α has to be sufficiently big to make the algorithm improve the diverseness of solution vectors. But, these values must be reduced in the last generations to get a superior and fine-tuned solution vectors. The values of p_a and α are varied in accordance with the number of generations and defined, where NI and gnrefer to the number of overall iterations and the present iteration, correspondingly [28].

$$P_a(gn) = P_{amax} - \frac{gn}{NI}(P_{amax} - P_{amin}) \quad (20)$$

$$\alpha(gn) = \alpha_{max} exp(c.gn) \quad (21)$$

$$c = \frac{1}{NI} Ln\left(\frac{\alpha_{min}}{\alpha_{max}}\right) \quad (22)$$

ICSA is a very successful in handling global optimisation since it is capable towards preserve stability among local and global random walks by means of switching limit. The parameters of the cuckoo are increased via the use of the gene values. This increases the classification results than the normal Cuckoo Search Algorithm (CSA).

4. RESULTS AND DISCUSSION

This section explains about the overall performance comparison analysis of the proposed technical work for the RA gene samples dataset. Factorsresponsible for radiographic severity of Rheumatoid Arthritis (RA) in African-Americans are interpreted in a poor light. This research work attempts to investigate genes whose expression in Peripheral Blood Mononuclear Cells (PBMCs) is associated with the radiographic severity of RA. 20 control samples (from individuals not affected with RA) were then compared with 10 early mild, 10 early severe, 10 late mild, and 10 late severe RA samples. All of these samples were acquired from African-Americans. Out of this 20 control samples, a total of 300 gene expression array were utilized in this analysis. This examination uses RA gene samples and metrics including precision, recall, accuracy and f-measure. In predictive analytics, a confusion matrix (Table 1), uses a Table having two rows and two columns associated with the number of False Positive(FP), False Negative(FN), True Positive(TP), and True Negative(TN).

Table 1: Confusion Matrix

		Predicted	
		Positive	**Negative**
Actual	**Positive**	TP	FP
	Negative	FN	TN

- TP is called as the amount of right predictions made that an instance is positive.
- FN refers to the amount of incorrect predictions made that an instance is negative.
- FP indicates the amount of incorrect predictions made that an instance is positive.
- TN stands for the amount of correct predictions made that an instance is negative.

Precision (P) is known as the ratio of the expected positive cases, which were right, as stated by expression (23)

$$P= TP/ (TP+FP) \quad (23)$$

The Recall or True Positive Rate (TPR) refers to the proportion of positive cases, which were correctly obtained, defined by equation (24)

$$TPR= TP/ (TP+FN) \quad (24)$$

The F-measure or balanced F-score (F1 score) is called as the harmonic mean of precision and recall. Through its multiplication with the constant 2, it makes the score as 1 when recall and precision are 1 in accordance with the equation (25)

$$F\text{-score}= 2TP/(2TP+FP+FN) \quad (25)$$

The accuracy is measured as the proportion of the total number of predictions, which were correct. It is expressed by equation (26)

$$Accuracy= (TP+TN)/ (TP+TN+FP+FN) \quad (26)$$

Table 2 provides the overall results of the techniques along with the metrics of performance evaluation.

Table 2: Performance Comparison Metrics vs. RA ClassificationTechniques

Techniques	Metrics			
	Precision (%)	**Recall (%)**	**F-measure (%)**	**Accuracy (%)**
SVM	77.18	78.20	79.50	66.67
FCM-PSO	84.44	82.61	83.52	75
DNCM-ICSA	94.02	93.1	92.12	89.12

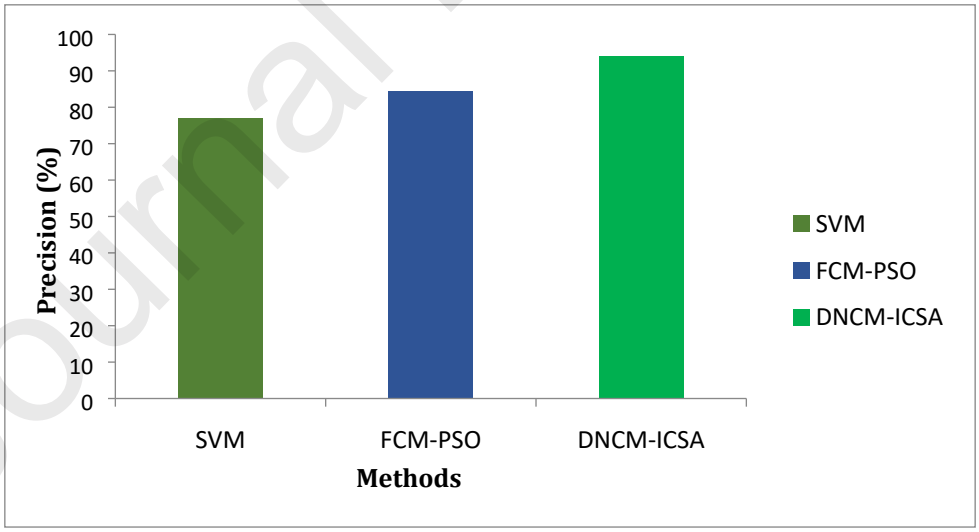

Figure 2: Precision Results Evaluation of RA Classification Methods

Figure 2 illustrates the performance results achieved of precision metrics corresponding to three classifiers such as the newly introducedSVM, FCM-PSO and DNCM-ICSA. The results show that the novelDNCM-ICSA classifier yields much better precision value of 94.02%, in comparison with the precision values ofSVM, FCM-PSOyielding 77.18%, 84.44% correspondingly.

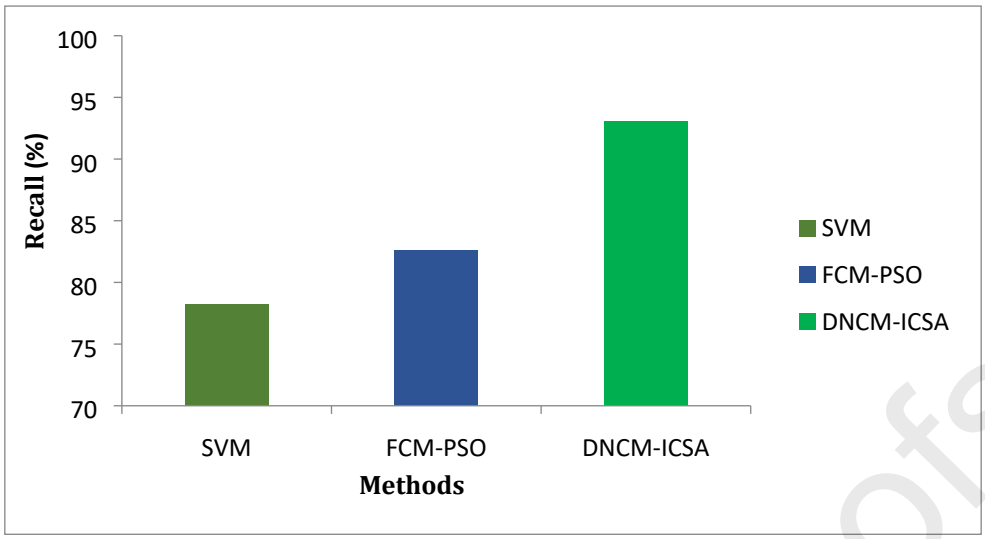

Figure 3: Recall Results Evaluation of RA Classification Methods

Figure 3 illustrates the recall results metrics corresponding to three classifiers such as the novelSVM, FCM-PSO and DNCM-ICSA. As observed from the figure 3 it is revealed that the novelDNCM-ICSA classifier yields much better recall value of 93.1 %, whereas the available techniques likeSVM, FCM-PSOyield much lesser recall value of 78.20%, 82.61% correspondingly.

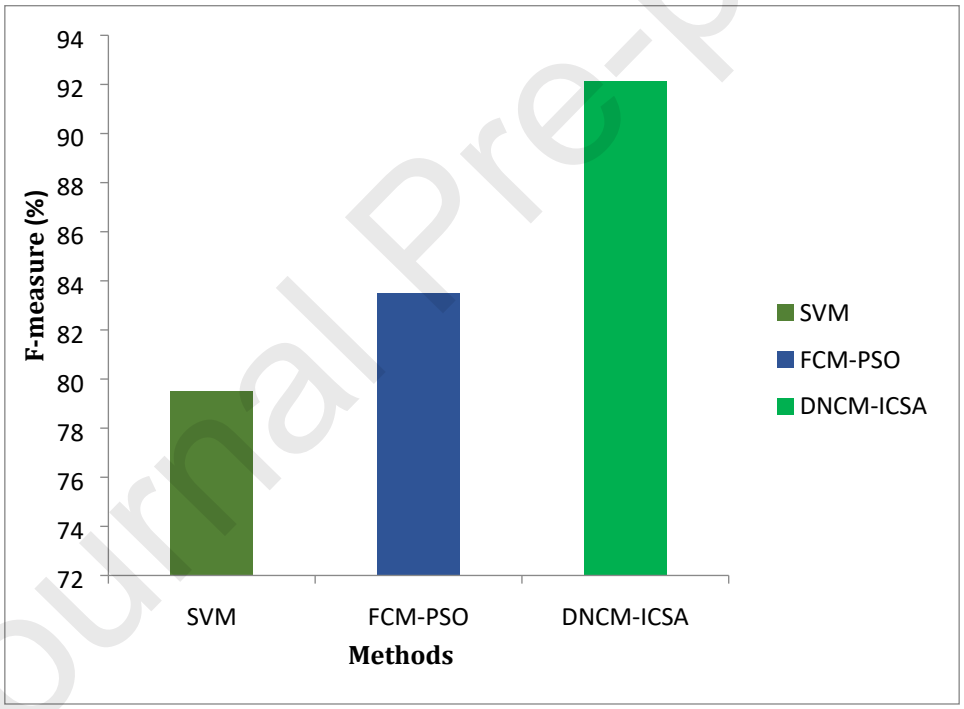

Figure 4: F-measure Results Evaluation of RA Classification Methods

Figure 4 demonstrates the results of the F-measure comparison analysis of the three classification techniques, which include SVM, FCM-PSO and DNCM-ICSA. The results reveal that the novel DNCM-ICSA classifier yield much better f-measure results of 92.12%, while the other available techniques like SVM, FCM-PSO yields 79.50% and 83.52% correspondingly.

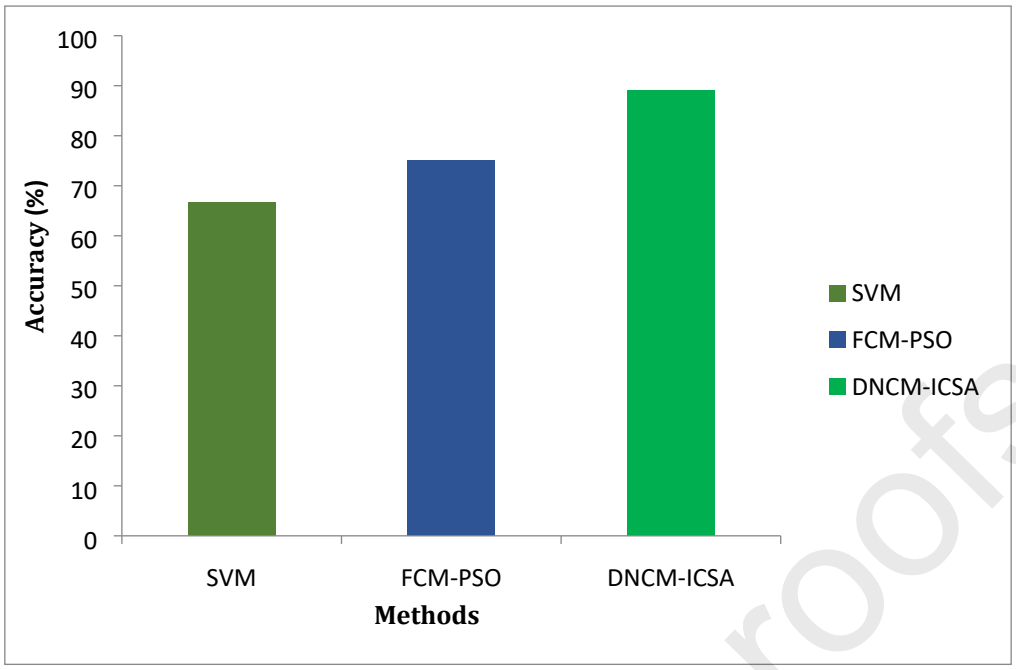

Figure 5: Accuracy Results Evaluation of RA Classification Methods

Figure 5 illustrates the results of the accuracy comparison analysis of the three classification techniques. The results show that the novel DNCM-ICSA classifier yields much better results of accuracy of 89.12%, while other available techniques like SVM, FCM-PSO yields 66.67% and 75% correspondingly.

4.1 Comparison of Cuckoo & Improved Cuckoo

The parameters and used in the CS assist the algorithm in finding both globally and locally extended solutions, correspondingly. The parameters and are quite significant parameters used in modifying the solution vectors, and can find huge application in making adjustments to the convergence rate of the algorithm. The classical CS algorithm makes use of constant value for both and α. These values get defined during the initialization stage and cannot be modified while newer generations spring up. The primary disadvantage of this technique shows up in the number of iterations to get an optimized solution. In case the value of is a small one and that of is big, and then the algorithm performance will be not so good and result in the number of iterations substantially increasing. In case the value of is big and that of is small, then the convergence speed becomes high, however it may be incapable of finding the best cluster. One primary variance between the ICS and CS lies in the way and α are adjusted. In order to boost the performance achieved of the CS algorithm and get over the disadvantages with static values of and α , the ICS algorithm makes use of variables and α. During the prime initial generations, the values of and α has to be sufficiently big to make the algorithm improve the diverseness of solution vectors. But, these values must be reduced in the last generations to get a superior and fine-tuned solution vectors. The values of and α are varied in accordance with the number of generations and defined, where and refer to the number of overall iterations and the present iteration, correspondingly. Comparison have been also done in the results section also in the figure 2-5 and table 2.

5. CONCULSION AND FUTURE WORK

It is quite evident that Rheumatoid Arthritis disease is a disease, which reduces the survival rate, leads toa higher level of disability and reduces the quality of human life. Dynamic Neutrosophic Cognitive Map with Improved Cuckoo Search Algorithm (DNCM-ICSA) with ensemble classifier is introduced for obtaining the gene expression profile that differentiates the patients affected with RA from significant control subjects. In this technical work, there are four important steps, which are involved, which include data preprocessing, feature selection, prediction and classification. During the first phase, gene selection refers to the procedure of preprocessing the input dataset with the objective of assessing theavailable attributes with the intent that only the genes relevant information are saved and irrelevant gene information are removed, in the next step, the gene selection is carried out through the T-test, chi-squared test, relief-F and Minimum Redundancy Maximum Relevance (mRMR). In the third step, disease prediction employing ensemble mechanism integrates the process of the Adaptive Neuro Fuzzy Inference System (ANFIS) and Deep Neural Networks (DNNs). At last, gene classification is performed by DNCM-ICSA algorithm. Prediction model and classification can be utilized for the detection of RA disease with E-GEOD-64707 dataset during the process of diagnosis. The resultant outcome

of the novel classifier are measured in terms of the metrics like precision, recall, F-measure and accuracy. At last,DNCM-ICSA classifier attains a much better accuracy value of89.12%, while other available techniques like SVM, FCM-PSO yields 66.67% and 75% correspondingly. As the work intended for the future, the following are proposed (i) design of novel hybrid classification techniques to boost the classifier accuracy and optimization of the computational efficacy of classifier, (ii)the prediction performance is increased by bringing in the hybrid feature selection approaches like filter-wrapper-embedded algorithm.

REFERENCES

[1] B. Heidari, Rheumatoid Arthritis: Early diagnosis and treatment outcomes, Caspian journal of internal medicine, Vol.2, No.1, Pp.161-170, 2011.

[2] Jr J.T. Birch and S. Bhattacharya, Emerging trends in diagnosis and treatment of rheumatoid arthritis, Prim Care, Vol.37, No.4, Pp.779–92, 2010.

[3] L. Gossec, C. Combescure, N. Rincheval, A. Saraux, B. Combe and M. Dougados, Relative clinical influence of clinical, laboratory, and radiological investigations in early arthritis on the diagnosis of rheumatoid arthritis, Data from the French Early Arthritis Cohort ESPOIR, The Journal of rheumatology, Vol.37, No.12, Pp.2486-2492, 2010.

[4] Q. Guo, Y. Wang, D. Xu, J. Nossent, N.J. Pavlos and J. Xu, Rheumatoid arthritis: pathological mechanisms and modern pharmacologic therapies, Bone research, Vol.6, No.1, Pp.1-14, 2018.

[5] C.Y. Chin, M.Y. Weng, T.C. Lin, S.Y. Cheng, Y.H.K. Yang and V.S. Tseng, Mining disease risk patterns from nationwide clinical databases for the assessment of early rheumatoid arthritis risk, PloS one, 2015.

[6] Z. Bedran, C. Quiroz, J. Rosa, L.J. Catoggio and E.R. Soriano, Validation of a prediction rule for the diagnosis of rheumatoid arthritis in patients with recent onset undifferentiated arthritis, International journal of rheumatology, 2013.

[7] A.N. Burska, K.Roget, M. Blits, L.S. Gomez, F. Van De Loo, L.D. Hazelwood, C.L. Verweij, A. Rowe, G.N. Goulielmos, L.G.M. Van Baarsen and F. Ponchel, Gene expression analysis in RA: towards personalized medicine, The pharmacogenomics journal, Vol.14, No.2, Pp.93-106, 2014.

[8] D. Aletaha, T. Neogi, A. Silman, J. Funovits, D. Felson, C. Bingham, N. Birnbaum, G. Burmester, V. Bykerk, M. Cohen, B. Combe, K. Costenbader, M. Dougados, P. Emery, G. Ferraccioli, J. Hazes, K. Hobbs, T. Huizinga, A. Kavanaugh, J. Kay, T. Kvien, T. Laing, P. Mease, H. Ménard, L. Moreland, R. Naden, T. Pincus, J. Smolen, E. Stanislawska-Biernat, D. Symmons, P. Tak, K. Upchurch, J. Vencovsky, F. Wolfe and G. Hawker, Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism Collaborative Initiative, Ann. Rheum. Dis., Vol.69, No.9, Pp.1580–1588, 2010.

[9] M.D. Mjaavatten and V.P. Bykerk, Early rheumatoid arthritis: the performance of the 2010 ACR/EULAR criteria for diagnosing RA, Best Pract. Res. Clin. Rheumatol, Vol.27, Pp.451-466, 2013.

[10] S. Whalen and G. Pandey, A comparative analysis of ensemble classifiers: case studies in genomics, IEEE 13th International Conference on Data Mining, Pp.807-816, 2013.

[11] B.V. Sumana and T. Santhanam, An empirical comparison of ensemble and hybrid classification, Proc. Processing and VLSI, Pp.463-470, 2014.

[12] S. Shanmugam and J. Preethi, Design of Rheumatoid Arthritis Predictor Model Using Machine Learning Algorithms, Cognitive Science and Artificial Intelligence, Pp.67-77, 2018.

[13] C. Lin, Y. Zou, J. Qin, X. Liu, Y. Jiang, C. Ke and Q. Zou, Hierarchical classification of protein folds using a novel ensemble classifier, PloS one, Vol.8, No.2, 2013.

[14] H. Sharon, I. Elamvazuthi, C.K. Lu, S. Parasuraman and E. Natarajan, Classification of Rheumatoid Arthritis using Machine Learning Algorithms, IEEE Student Conference on Research and Development (SCOReD), Pp. 245-250, 2019.

[15] Z. Shiezadeh, H. Sajedi and E. Aflakie, Diagnosis of rheumatoid arthritis using an ensemble learning approach, ICAITA, SAI, CDKP, Signal, Pp.139-148, 2015.

[16] S. Bardhan and M.K. Bhowmik, 2-Stage classification of knee joint thermograms for rheumatoid arthritis prediction in subclinical inflammation, Australasian physical & engineering sciences in medicine, Vol.42, No.1, Pp.259-277, 2019.

[17] S. Shanmugam and J. Preethi, A Study of Early Prediction and Classification of Arthritis Disease using Soft Computing Techniques, International Journal for Research in Engineering Application & Management (IJREAM), Vol.03, No.05, Pp.35-47, 2017.

[18] J.L. Salmeron, S.A. Rahimi, A.M. Navali and A. Sadeghpour, Medical diagnosis of Rheumatoid Arthritis using data driven PSO–FCM with scarce datasets, Neurocomputing, Vol.232, Pp.104-112, 2017.

[19] M. Kourilovitch, C. Galarza-Maldonado and E. Ortiz-Prado, Diagnosis and classification of rheumatoid arthritis, Journal of autoimmunity, Vol.48, Pp.26-30, 2014.

[20] S. Singh, A. Kumar, K. Panneerselvam and J.J. Vennila, Diagnosis of arthritis through fuzzy inference system, Journal of medical systems, Vol.36, No.3, Pp.1459-1468, 2012.

[21] S.R. Pandey, J. Ma and C.H. Lai, Development of decision support system for the diagnosis of arthritis pain for rheumatic fever patients: based on the fuzzy approach, Journal of Algorithms & Computational Technology, Vol.9, No.3, Pp.265-290, 2015.
[22] Y.K. Jain and S.K. Bhandare, Min max normalization based data perturbation method for privacy protection, International Journal of Computer & Communication Technology, Vol.2, No.8, Pp.45-50, 2011.
[23] N. Ziasabounchi and I. Askerzade, ANFIS based classification model for heart disease prediction, International Journal of Electrical & Computer Sciences IJECS-IJENS, Pp.7-12, 2014.
[24] K.H. Miaoa and J.H. Miaoa, Coronary Heart Disease Diagnosis using Deep Neural Networks, International Journal of Advanced Computer Science and Applications, Vol.9, No.10, Pp.1-8, 2018.
[25] J. Kukačka, V. Golkov and D. Cremers, Regularization for deep learning: a taxonomy, Artificial Intelligence, Pp.1-27, 2017.
[26] G. Seni and J.F. Elder, Ensemble methods in data mining: improving accuracy through combining predictions, Synthesis lectures on data mining and knowledge discovery, 2010.
[27] S.I. Tusiy, M. NasifShawkat, A. Ahmed, B. Panday and N. Sakib, Comparative analysis of improved cuckoo search (ICS) algorithm and artificial bee colony (ABC) algorithm on continuous optimization problems, International Journal of Advanced Research in Artificial Intelligence, Vol.4, No.2, Pp.14-19, 2015.
[28] J. Wang, B. Zhou and S. Zhou, An improved cuckoo search optimization algorithm for the problem of chaotic systems parameter estimation, Computational intelligence and neuroscience, 2016.
[29] Ki-Jo Kim and Ilias Tagkopoulos, Application Of Machine Learning In Rheumatic Disease Research, Korean J Intern Med., 2019 Jul; 34(4): 708–722.
[30]. Cristina Regueiro, Lorena Rodríguez-Martínez, Laura Nuño, Ana M. Ortiz, Alejandro Villalba, Dora Pascual-Salcedo, Ana Martínez-Feito, Isidoro González-Alvaro, Alejandro Balsa, and Antonio Gonzalez, Improved RA Classification Among Early Arthritis Patients With the Concordant Presence of Three RA Autoantibodies: Analysis in Two Early Arthritis Clinics, Arthritis Research & Therapy, 2019 Dec 11;21(1):280. doi: 10.1186/s13075-019-2079-4.
[31] S. Shanmugam & J. Preethi, Improved feature selection and classification for rheumatoid arthritis disease using weighted decision tree approach (REACT), The Journal of Supercomputing, 75, pages5507–5519(2019).