ISSN: 0974-5823 Vol. 7 No. 4 April, 2022

International Journal of Mechanical Engineering

ENSEMBLE MACHINE LEARNINGAND HYBRID NEUTROSOPHIC COGNITIVE MAPS BASED FEATURE SELECTION FOR RICE DISEASE PREDICTION

T.P.Senthilkumar, Assistant Professor of Computer Science, Gobi Arts & Science College, Gobichettipalayam.

Dr.P.Prabhusundhar, Assistant Professor of Computer Science, Gobi Arts & Science College,Gobichettipalayam.

Abstract: Rice is a major crop that is grown all over the world. Several diseases and pests attacks reduce its output quality. The diseases which affect Rice include Leaf Blasts, Brown Spots, rotten Stems, Bakanae, and Yellow Dwarfs which affect different parts of the crop. Farmers can maintain both the quality and quantity of their crop by detecting disease early and implementing preventative measures. However, the existing method cannot perform well in high dimensional dataset where as the error are occurred in the feature selection process which will directly affect the classifier performance. So this research work is introduced an Ensemble machine learning and Hybrid Feature Selection methods for efficient prediction of rice diseases. Initially, the SMOTE based pre-processes are proposed for data normalizations. And then the Modified Feature weighted Fuzzy Clustering (MFWFC) based segmentation is proposed for efficient segmentation. Then the feature extraction is carried on the EICAs (Enhanced Independent Component Analyses) are developed for increasing classifier performancea. The feature selection is done by using HNCMs (Hybrid Neutrosophic Cognitive Maps). And finally the ECMs (Ensemble Classification Models) are used to improve prediction performances. Here the MGNNs (Modified Granular Neural Networks), Extreme Gradient Boost Classifiers and LRs (Logistic Regressions). Experimental results demonstrate usefulness of the proposed model with higher performances when compared to existing models.

Keywords: Rice Disease, SMOTE, Modified Feature weighted Fuzzy Clustering (MFWFC), Enhanced Independent Component Analysis (EICA), Hybrid Neutrosophic Cognitive Maps (HNCM), Classification Model (ECM), Modified Granular Neural Network (MGNN), Extreme Gradient Boost Classifier and Logistic Regression (LR).

1. Introduction

Agriculture is vital to the Indian economy and is second in rice productions. Rice is cropped in literally every state including Tamil Nadu, West Bengal, Punjab, Uttar Pradesh, Assam, and Bihar, to name a few [1]. The agricultural industry contributes roughly 19.9% of the overall gross domestic output. In India, rice is one of the most widely consumed grains. Diseases impair the development and quality of rice plants, lowering the profitability of the cultivation. Different illnesses can affect specific rice crops, making it difficult for farmers to identify them due to their limited expertise gathered through experience. Automatic data processing expert is used for this accuracy and early detection of plant diseases diagnosis. As a result, the development of a healthy and prosperous crop is possible [2]. In agriculture, computer vision technology has been used to estimate crop yields, diagnose crop nutritional deficiencies, estimate crop geometric sizes, and recognise weeds for decades. Image processing, pattern recognitions, SVMs (support vector machines), and hyperspectral detections have been used in computer vision to diagnose diseased crops or plants . Cluster analysis was utilised to distinguish healthy tomatoes from diseased in multispectral remotely sensed images of tomato farms [3]. SVMs extracted form and textural aspects of rice bacterial leaf blasts and leaf, sheath blights. The study used GAs (Genetic Algorithms) in combination SVMs with to categorise rice brown spots, bacterial blights, and leaf blasts and thus recognize sick leaves in plants. RGB (Red Green Blue) values of afflicted parts followed by application of NBs (Naive Bayes) in classifications. Tomato's mosaic disease and wheat leaf rusts were identified utilising infrared thermal imaging where crop's temperature information were retrieved. Though many existing algorithms detect crop diseases with decent accuracies, they are heavily dependent on manual feature extractions making the generalizations of these extractions difficult [4]. Furthermore, certain approaches need the use of specialised equipment, which is not always easily available to users. All of these flaws make it difficult to use these technologies to diagnose agricultural diseases.

Filtering, wrappers, and embedding approaches are the three types of feature selection strategies. Filter approaches look at the data's inherent traits, and feature selection is independent of the classifier [5]. As part of the selection process, wrapper models require improving a predictor. Although they produce superior results, filter approaches are often less computationally costly than wrappers. The embedded technique incorporates the selection procedure within the classifier's learning process. There are obstacles in identifying effective features by algorithms due to noises or redundancy or irrelevant information present in the data added to their higher dimensionalities [6]. This makes selection of efficient feature selection techniques challenging. Feature selection's performances are reliant on their success learning of data. Instability of feature selection methods were reviewed in the overview of feature selection strategies. Multiple selection strategies including filters, wrappers, and embedding techniques were

examined in terms of feature selections. Feature selections using wrappers or ensembles employ their own learning processes but may not perform well with other classifiers. Filter-based approaches have less computational complexity than embedded and wrapper-based techniques. Wrapper-based approaches have a considerable risk of overfitting due to their complexity [7]. Filter-based techniques generate stabilized sets of features which resist data over fits. More generic strategies for feature selections are ensembles. A commonly employed filter technique that analyzes intrinsic properties of inputs without classifiers on mind is used for feature selections in this work called Neutrosophic logic. The maps generated by NCMs are neutral between true and false and useful filters for selection of features [8] where lesser number of characteristics characterise the complete data. Many filters and wrappers including MCFSs (multi-cluster feature selections), Relieff, LSs (Laplacian Scores), Inf-FSs (infinite feature selections), ILFSs (infinite latent feature selections), local learning-based CFSs (clustering feature selections), and GAs are compared with this works suggested neutrosophic logic-based feature selections. These algorithms are intended to rank and choose relevant traits as feature subsets according to relevance or significance to applications. The focus of this research is on the hybrid feature selection method.

To circumvent disadvantages of crop disease diagnostic methods DLTs (Deep learning techniques) could be used. DLTs have become popular in recent years for image categorizations, object identifications, and content suggestions. In fact, DLTs have been employed by researchers to identify illnesses in a variety of crops [9]. To identify diseased wheat crops, researchers first used CNNs (convolution neural networks), Faster R-CNNs (Faster Recurrent CNNs) on sugar beet's leave imagers. Data Over fits occus while using MLTs (Machine Learning Techniques) based on single prediction models whereas ensemble learn using series of algorithms and integrate feasible predictions thus addressing this issue. Ensemble learning has been used to predict illness diagnosis, soybean production, protein binding hot spots, and wheat grain yield since the introduction of computer technology. Since, ensembles have been established as a strategy in earlier researches, they were used in this work to improve accuracy of sickness detections [10]. To summarise, DLTs have the potential to identify diseases in a wide range of crops with higher accuracies. Existing researches on use of DLTs to trace rice ailments concentrate on limited number of diseases amongst identified diseases in Rice crops which include rice leaf blasts, fake smuts, neck blasts, sheath blights, bacterial stripes, and brown spots [11]. However, existing methods do not perform well on high dimensional datasets where errors mainly stem from their feature selection processes impacting classifier's performances. Hence, this research work introduces Ensembles and Hybrid Feature Selection methods for efficient prediction of diseased rice crops. The goal of this study was to improve diseased rice crops detection accuracies, efficiencies, costs, and conveniences. The study's particular goals were to (1) construct ensemble learning models for diagnosing various types of rice illnesses, (2) evaluate proposed model's performances, and (3) implement the suggested diagnostic approaches.

The rest of the research work is organized as follows as, section 2 review the some of the recent technique for developing the machine learning based diseased rice crops prediction. section 3 presents the proposed methodology. section 4 provides the conclusion and its future work.

2. Literature Review

In recent years, many DLTs and MLTs constructed with different architectures have been proposed for diagnosing rice illnesses. This study reviews some of these recent studies for detecting diseased rice crops using advanced data mining techniques in this section.

Hsieh et al [12] trained their categorizations using Auto-Sklearn and NNs (neural networks). Their results of tests suggested that their proposed model was 72% accurate in classifications of worse RBD scenarios.In the instance of exacerbation, our model achieves an accuracy of 89 percent, demonstrating the usefulness of the suggested categorization approach. Das et al [13] aim of increasing agricultural outputs proposed automated feature engineering based on DLTs for predicting diseased rice leaves and thus eliminate of minimize disease propagations from infected regions. Damaged rice leaves were first identified and separated from the rest. The diseased images were submitted to CNNs. Their CNNs had four convolution, two completely connected and one softmax output layers. The reason for using CNNs were to infinite number of features without any bias using automated feature engineering and capture intricate non-linear interactions between features. The illnesses are then categorised using several classifiers after a dimension reduction approach is used to eliminate redundant characteristics. A total of 10,500 contaminated leaves were used to test the procedure. The experimental findings and a comparison analysis based on performance assessment demonstrate the efficacy of the suggested strategy and aid in the selection of the best classifier for rice leaf disease prediction. Bashir et al. [14] used an SVM (Support Vector Machine)-based technique to classify and assess three rice crop illnesses. The procedure is divided into two phases: training and illness prediction. Using a trained classifier, the method detects illness on the leaf. The suggested research study maximises the efficiency of SVM parameters (gamma, nu). The suggested method obtained 94.16 percent accuracy, 5.83 percent misclassification rate, 91.6 percent recall rate, and 90.9 percent precision, according to the findings. The comparative findings show that the proposed technique is effective and yields high accuracy percentage as compared to the other techniques. By adding image processing and collaborative elements, the results produced can aid in the creation of a successful software solution. This may make it easier for farmers and other organisations to make informed decisions in order to safeguard rice harvests from significant harm. In light of the outcomes of this study, the given approach might be a viable option for incorporating image processing techniques into KM (Knowledge Management) systems. Pinki et al. [15] suggested diagnosed Brown spots, Leaf blasts, and Bacterial blights in paddy leaves with insecticides and/or fertilisers recommended based on disease's severity. The impacted region of the paddy leaf picture is separated using K-means clustering. These illnesses are classified based on their visual contents (colour, texture, and form). SVMs classified paddy leaf diseases and subsequently predictive cures were offered with the aim of assisting agriculture-related individuals and organisations in take suitable measures to combat these illnesses..

Sharma et al [16] used CNNs categorize disease predictions in rice crops. Paddy diseased crops or plants are extremely dangerous and can seriously harm crops if not addressed early on. The suggested model would enhance decision making in the case of various illnesses in rice crops utilising CNNs for early disease predictions thus limiting yield losses to a minimum. Sabitha et al [17] used data on sick leaf sections extracted by threshold segmentation to construct fuzzy judgement algorithms for identifying rust severity levels in images. Six distinct colour and texture qualities were used to determine the results of these trials. Clustering of plant diseases was done using Fuzzy Algorithms where the study's results were more efficient than traditional methods, placing it as the best amongst feature extraction strategies and thus use them for diagnosing plant diseases from leaf images. To describe these capabilities, further disease categories or crop/disease classifications can be introduced. Karthick et al [18] implemented weight-based similarity techniques for type II full-spaces and sub-space ensembles, including random-k, fixed-k approaches. For clustering categorical datasets, clustered ensemble similarity frameworks were presented and cluster centroids were chosen using bio-inspired/swarm-based optimizations. DRHs (Division Reminder Hashes) determined cluster weights based on similarity assessments in this suggested method for eliminating null and duplicate values from datasets. AR (Adjusted Rand), NMI (Normalized Mutual Information), and CAs (Classification Accuracies) were the testing measures used in the study The proposed strategy's performance was evaluated using UCI (University of California Irvine) datasets, and the findings demonstrated that their proposed approach outperformed other Type-II and Type-III ensemble clustering. Das et al [19] built weighted networks with nodes representing retrieved attributes and weights denoting similarities between feature pairs. The study used Louvain community discovery approach for division of graphs. The most significant node in graph partitions were selected as the partition's highest-ranked node and the graphs were used to determine subset of characteristics. The importance of the proposed method was demonstrated by its successful predictions of diseased rice crops.

CNNs based deep learning introduced by Krishnamoorthy et al [20] has been succesfully invoked for handling computer vision issues such as picture classification, object segmentation, image analysis, and so on. InceptionResNetV2 is a form of CNN model used in our research to identify illnesses in rice leaf photos using a transfer learning technique. The suggested model's parameters were tuned for the classification job, and the resulting accuracy was 95.67%. Sengupta et al [21] used PSOs (Particle Swarm Optimizations) and ARM (Association Rule Mining) to build incremental rule-based categorization systems, since, rice disease characteristics change over a period of time due to changes in meteorological, biological, and geographical factors. The proposed incremental classifications were well suited for disease predictions on diseased rice crop datasets. The study's evaluations on simulated sick rice crop datasets and benchmark datasets in terms of accuracies showed better results when compared to existing classification methods. The procedures were further assessed using statistical measurements and statistical tests to determine its relevance and efficacy. Deng et al [22] proposed ensembles for efficient diagnosis of rice diseases where sub-models were integrated. The proposed ensembles namely DenseNet-121, SE-ResNet-50, and ResNeSt-50 were validated using a second set of images and on metrics of learning rates, precisions, recalls, and disease recognition accuracies. The study's ensembles minimised rice disease misdiagnosis and achieved an accuracy of 91% on six verieties of rice crop diseases.

From the above discussion, it is identified that the existing methods have some of the issues as well. And also it cannot perform well in high dimensional dataset where as the error are occurred in the feature selection process which will directly affect the classifier performance. So this research work is introduced an Ensemble machine learning and Hybrid Feature Selection methods for efficient prediction of diseased rice crops .

3. Proposed Methodology

This research work is introduced an Ensemble machine learning and Hybrid Feature Selection methods for efficient prediction of diseased rice crops . Initially, the SMOTE based pre-processes are proposed for data normalization. And then the MFWFC based segmentation is proposed for efficient segmentation. Then the feature extraction is carried on the EICAs are developed for increasing the classifier performance. The feature selection is done by using HNCMs . And finally the ECMs improve prediction performances. Here the MGNNs , Extreme Gradient Boost Classifier and LRs . The figure 1. shows the process of the proposed methodology.

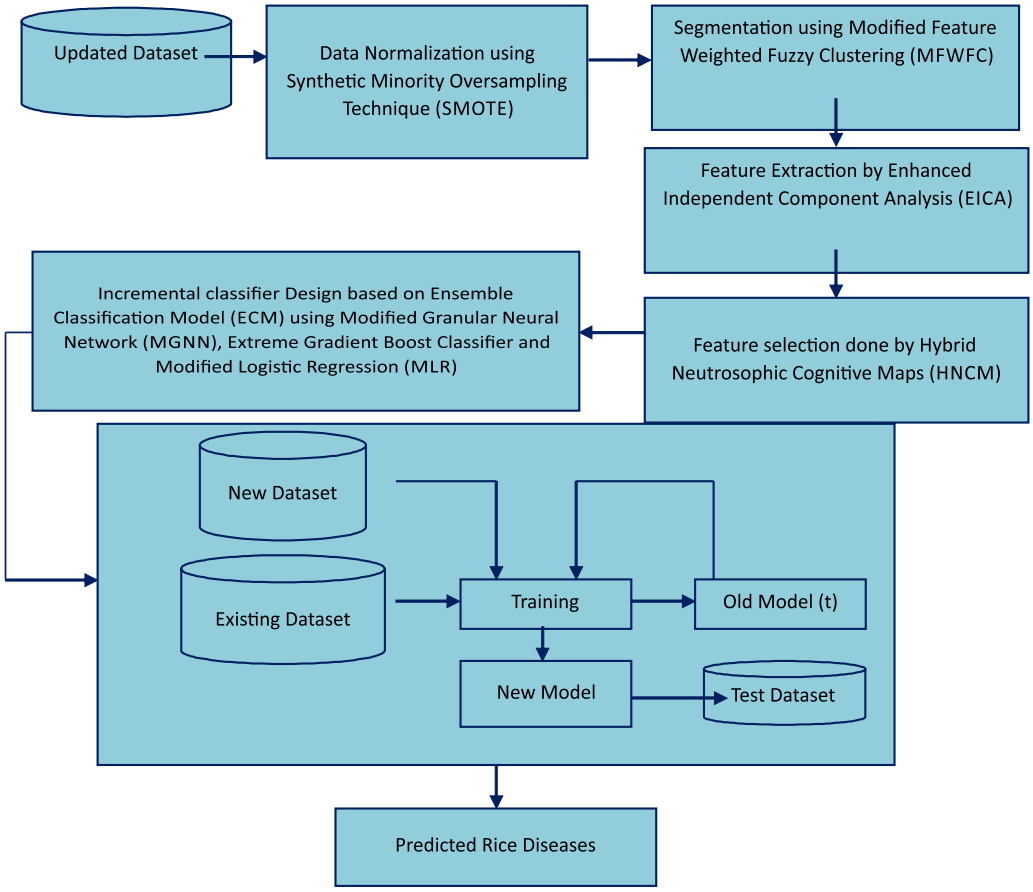


Figure 1. The overall process of the proposed methodology

3.1. Dynamic classifier for incremental data

Incremental learning represent ad hoc learning which occur as new training data enters over a period of time, as opposed to fixed static learning methods based on fixed training datasets. Incremental learning techniques are continuous learning procedures that are suited for software agent's learning where agents adapt to dynamic environments progressively [23]. This research work uses incremental classifiers with the aim of having the least feasible classification rules feasible. This strategy can successfully handle incremental data for optimum rules by modifying its present knowledge base to accommodate new data. The proposed incremental technique is depicted in Figure 1. Initially ECMs, train on existing dataset to establish optimum classification criterions. When new batches of data are received, incremental ECMs generate dynamic classifications using previous classifiers and new sets of data. The classification performances are evaluated on test datasets. As a consequence, the suggested ECMs analyze new datasets frequently basis and dynamically update their prior knowledge bases in reduced training times. The steps of the proposed ECMs with Incremental Classifications are as follows:

3.2. Dataset preparation

The datasets are handled as incremental datasets to show the approach using benchmark datasets. As a result, each dataset designated as a decision system DS is subdivided into DS1, DS2, and DS3 subsystems. DS1 is an old dataset, DS2 is a new or incremental dataset, and DS3 is a test dataset intended to assess the ECM algorithm's performance. Because the approach only works with discrete valued datasets, the continuous dataset must be discretized before it can be used..

3.3. Preprocessing using SMOTE

Data pre-processing is a common initial step before training and analysing the data using MLTs. Data used to train MLTs need to be formatted effectively for consistency and best possible outputs and where significant characteristics need to be included [24]. Pre-processing data for MLTs involve multiple processes, including normalizations. The purpose of these processes is to remove less relevant characteristics in data sets. In order to provide better predictions, most classification algorithms seek pure examples in learning and create boundaries of classes. Synthetic cases that are distant from boundaries are easier to classify than those that are close to the boundaries and poses considerable learning challenges for most classifiers. Based on these findings, this study uses an advanced technique called A-SMOTE for pre-processing imbalanced training sets and with the goal of characterising borderlines and creating cleaner synthetic samples from SMOTE generalisations. The proposed method is separated into two sections as detailed below:

First stage, used SMOTE techniques [25] to create synthetic instancesfollowing equation (1)

$$N = 2 * (r - z) + z \qquad (1)$$

Where N is the initial synthetic instance number (which is produced from scratch), r is the number of majority class samples, and z is the number of minority class samples.

In the second stage**,** synthetic samples similar to majority classes and synthetic instances near SMOTE's borders are eliminated. For cleaning provided data, segmentations are used, as outlined in the following sections.

3.4. MFWFC based segmentation

FCMs (Fuzzy C-means Clusters) are a type of unsupervised fuzzy clustering algorithm that is based on objective function optimization. Assuming $X = \{x_1, x_2, \ldots, x_n\}$ represents a collection of n pixels, x_j is the data's eigenvalue, c represents the count of categories, and $V = \{v_1, v_2, \ldots, v_c\}$ are collections of cluster centers and J_m objective function satisfying the constraint:

$$\sum_{i=1}^{C} u_{ij} = 1\left(u_{ij} \in [0,1]\right) \quad (2)$$

$$J_m = \sum_{i=1}^{C} \sum_{j=1}^{n} u_{ij}^m d^2\left(x_i, v_j\right) \quad (3)$$

Where u_{ij} is j^{th} sample's membership in the i^{th} category; m represents weighed index, and generally $m = 2$; and $d^2(x_j, v) = \|x_j - v_i\|^2$ stands for Euclidean distances of x_j from v_i (cluster centers). Objective function J is the sum of the squares of weighed distances between data points and cluster centres. Clustering effects are greatest in data points located near cluster centres and FCM membership functions only consider similarity measures between intensity features and cluster centres, not their spatial features. According to this study, attribute features and spatial characteristics are assigned distinct weights independently. using symmetric distribution factors, and then fused in terms of particular criteria to overcome the constraint.

- **Symmetric Distribution factor**

Assume n features are conditionally independent, then applying symmetric distributions to observed probability values of features and computing their products after normalisations are based on:

$$p(z^1 \ldots z^n | x) = \prod_1^n \left(\frac{p\left(z^i | x\right) + \beta^i U(x)}{1 + \beta^i} \right) \quad (4)$$

where β^i is the uncertainty of i, $U(x)$is the discrete Steady state distribution, n is the number of features.

Next, the corresponding weight values are assigned to different features using the Symmetric Distribution factor δ_{DF} calculation formula is as follows:

$$\delta_{DF} = b.\left(-k.rand(\quad) + tan\left(\frac{\pi}{4} - \frac{\pi l}{4.Max_{iteration}} \right) \right) \quad (5)$$

where l is the current iteration count, Max_Iteration represents maximum iterations, b stands for proportion coefficient used for avoiding imbalances. The disturbance deviation factor is represented by k and the rand function is used to disrupt falling δ_{DF}values. The updated counts of leaders in iterations is $\_\delta_{DF} \cdot N$, and the number of followers equals $1 - \delta_{DF}$. From the above model, the efficient improvement is done in segmentation process for the given data. the below section describes the feature extraction process.

3.5. Feature Extraction using EICAs

ICAs (Independent Component Analyses) are relatively new methods which aim to represent non-Gaussian data linearly so that components display statistical independence to maximum feasible extents. These formats determine fundamental structures of feature extractions as well as signal separations [26]. The Notion of independence can be demonstrated using 2 random scalars namely y_1 and y_2. These variables are considered independent when no information of y_1 has anything to do with information in y_2. Independence was found to be true for the variables s_1,s_2, though there were some dependencies' between mixed variables x_1,x_2.

Formally, density probabilities can also be used to define independence. Assuming (y_1,y_2) represent combined PDFs (Probability Density Functions) of y_1 and y_2 and if $p_1(y_1)$ represents marginal PDFs of y_1 or PDFs while considering y_1 separately:

$$p_1(y_1) = \int p(y_1, y_2) dy_2 \quad (6)$$

Which is also the same for y_2, the random variables y_1 and y_2 are can be called independent iff their joint PDFs are factorizable according to Equation (7).

$$p(y_1, y_2) = p_1(y_1) p_2(y_2). \quad (7)$$

This definition generally projects joint densities as product of n terms for n random variables, thus assisting in finding out significant properties of independent random variables. This is done using functions, h_1 and h_2 where:

$$E\{h_1(y_1) h_2(y_2)\} = E\{h_1(y_1)\} E\{h_2(y_2)\}. \quad (8)$$

Estimation in ICAs are borrowed from information theory for reducing mutual information which is similar to the notion of obtaining majority of non-Gaussian directions. Mutual Information Using differential entropy concepsts, mutual information I between m (scalar) random variables can be computed, y_i, i = 1...m is defined as below,

$$I(y_1, y_2, \ldots\ldots, y_m) = \sum_{i=1}^{m} H(y_i) - H(y). \quad (9)$$

Mutual information are common metrics used to assess degrees of interdependences between random variables and in practises, a widely employed measure of independence are common divergences between joint densities f(y) and products of their marginal densities whici is always positive, and zero when variables are statistically independent. As a result, mutual information includes entire dependence of the variables, not just the covariances.

To comprehend mutual information, entropy may be interpreted as code length. When the encoding is done individually, the words H(yi) provide the code lengths for y_i, but H(y) is coded as a random vector, suggesting that all components have same code values As a result, mutual information reveals how much of code lengths are saved when complete vectors are encoded instead of separate component encodings. In general, acceptable codes are obtained by completely encoding vectors. If y_i is self-contained, they have no knowledge of one another, and variables can be encoded individually without increasing coding lengths. One significant characteristics of mutual information for invertible linear transformations y = Wx are:

$$I(y_1, y_2, \dots \dots, y_n) = \sum_i H(y_i) - H(x) - log \vee detW \vee .(10)$$

Also, for y_i of unit variance, entropy and neg entropy are different just by a constant, and the sign. Therefore, the following expression,

$$I(y_1, y_2, \dots \dots, y_n) = C - \sum_i J(y_i)(11)$$

where C refers to a constant, which has no dependency on W which represents basic associations between negentropy and mutual information.

The bases are changed to increase generalisation capabilities of classifiers. To decrease intra-class variations, environmental factors are considered, however, different implementations of ICAs representing source distributions change placements of basis vectors, resulting in features being measures of invariance. ICAs seek ortho-normal rotations in whitened spaces, which provide statistical independences and preferences for ICAs. There are many efficient techniques for lowering dimensions, but ICAs based on Enhanced Probability Distributions are commonly utilised. They are fixed iterations that aid in maximising single and multiple independent components in probability objective functions.

EPDs (Enhanced Probability Distributions) based ICAs are adaptive maximum likelihood estimates that aid in modelling the source distributions by taking skewness into account. The modelling method generates a meaningful relationship between real estimators and theoretical measures of independences. EPD score functions are used as objective functions in ICAs that need to be maximised. For maximising, natural gradient and fixed-point methods based on EPD frameworks are presented. The PDs (Probability Distributions) are expressed by inverse distribution functions:

$$F^{-1}(p) = \lambda_1 + (p\lambda_3 - (1-p)\lambda_4)/\lambda_2 \qquad (12)$$

Where 1>= p >= 0 and parameters used for distributions are λ_1, λ_2, λ_3, and λ_4. PDs are true when 0 <= λ_2 / (λ_3 pλ_{3-1} – λ_4 (1-p) λ_{4-1}). Correlations λ_1, λ_2, λ_3,$\lambda_4$and moments α_1, α_2, α_3, and α_4 are defined using 4 non-linear equations resulting in numerical values. Density functions of EPDs do not exist in closed forms and score functions are obtained from inverse distribution functions of $p = F(y)$ where EPD's distribution functions are represented by $F(y)$. Score functional observations are computed numerically finding solutions for values of p and applying score functions ϕ(p).

$$\hat{\alpha}_1 = \acute{x} = \sum_{i=1}^{n} x_i/n, \qquad (13)$$

$$\hat{\alpha}_2 = \hat{\sigma}^2 = \sum_{i=1}^{n} (x_i - \acute{x})^2/n, \qquad (14)$$

$$(x_i - \acute{x})^3/n, (\hat{\sigma}^3), \qquad (15)$$

$$\hat{\alpha}_3 = \sum_{i=1}^{n} (x_i - \acute{x})^3/n (\hat{\sigma}^3).$$

$$(x_i - \acute{x})^4/n (\hat{\sigma}^4). \qquad (16)$$

$$\hat{\alpha}_4 = \sum_{i=1}^{n}$$

The correlations between parameters and moments found numerically using equations are complex computations. Hence, λ_1, λ_2, λ_3, and λ_4 are considered as functions of α_3 and α_4 for standardized data where α_1=0 and α_2 =1.

The fundamental source distributions are estimated using marginal distributions that have been adapted to the EPD family using the technique of moments as stated. Because the EPD's density function does not exist in closed forms, scores are functionally generated from inverse distributions.

$$p = F(y)(17)$$

where F(y) refers to the distribution function of an EPD, and the following formula is got for the score function:

$$\lambda$$

$$\lambda_2(1-p)^{\lambda_4-2}\frac{(\quad|4-1)\lambda_4}{\left(p^{\lambda_3-1}\lambda_3+(1-p)^{\lambda_4-1}\lambda_4\right)^2}. \quad (18)$$

$$\varphi(p)=\frac{-\lambda_2 p^{\lambda_3-2}(\lambda_3-1)\lambda_3}{\left(p^{\lambda_3-1}\lambda_3+(1-p)^{\lambda_4-1}\lambda_4\right)^2}+$$

The original algorithm that optimizes the criterion derived where maximum likelihood contrasts are used.

$$W_{k+1}=W_k+\eta(I-\varphi(y)y^T)W_k \quad (19)$$

whereηrefers to the learning rate and fixed point algorithm,

$$W_{k+1}=W_k+D\big(E\{\varphi(y)y^T\}-diag(E\{\varphi(y_i)y_i\})\big)W_k \quad (20)$$

where D = diag$\left(\frac{1}{E\{\varphi(y_i)y_i\}-E\{\varphi'(y_i)\}}\right)$.

A process for the EPD-ICAs may be provided as below: This process is repeated until convergence criteria are satisfied.

1. Compute the third and fourth sample moments α_3 and α_4 for current data $y_k=W_k x$and choose the EPD if $\alpha_4>2.2+2*\alpha_3^2$.
2. Estimate the parameters for EPD using the method of moments and compute scores$\varphi(y_k)$.
3. Compute the demixing matrix W_{k+1}.

The features are extracted in a intelligent manner with the help of Enhanced Independent Component Analysis. And the next process is feature selection which is explained in the below section.

3.6. Feature Selection using HNCMs

Hybrid NCMs, a filter-based methodology, is employed to choose relevant and most dominant characteristics for rice illness categorization in the suggested method [27]. To improve the accuracy rate, the Modified Bat method is used with the NCM model..

NLs (Neutrosophic logics) combine fuzzy logics, para-consistent logics, intuitionistic logics. In NLs, all logical variables are represented as letters namely T (truths), I (indeterminacies) and F (falsehoods). The variables are represented by the union and intersection of unique subsets, intervals, single-finite components that can be infinite/finite, continuous/discontinuous, and any real sub-unitary subsets. The human mind's representations are analogous to NLs which aim to reflect imprecision caused by uncertainties or ambiguities of varied observers due to limited information, resulting in T, I, F being subsets instead of single elements. If map's edge values in NLs are from the set {0,1, I}, where 1 is for truth, 0 for falsehood, and I for indeterminacy then NCMs can be formed from maps of NLs.

Table 1 Weights of nodes in NCM

Degree of member-ship	Effects
− 1	Increase in C_i results in decrease in C_j
0	C_i have no effect on C_j
1	Increase in C_i causes increase in C_j
I	The relation between C_i and C_j is indeterminate

The fundamental reason NLs are important for leaf disease detection analyses are that I exposes indeterminacy and (ii) suggesting that good traits in one system may be misleading in another. The dichotomy between absolute and relative truths is the major difference between intuitionistic fuzzy logics and NLs. Logical assertions are represented in three-dimensional spaces in NLs using F, T, and I. NCMs can be defined as follows:

Definition 1:NCMs are directed graphs that represent causal relationships between features.

Definition 2: Nodes in NCMs are fuzzy nodes when fuzzy sets for its nodes.

Definition 3: Each node in the graph is considered as a feature. Weight is assigned to each directed edge between nodes (Ci and Cj). The weight value lies in the set {− 1, 0, 1, I}. Table 1 shows the weights of nodes in NCM. The following steps are described in the below,

Step 1: The patient assigns linguistic variables (small, medium, and high) to the symptoms, which are then transformed into crisp numbers.

Step 2: Then, in a range of 0 to 1, these crisp values are fuzzified.

Step 3. An starting vector is produced using the specified symptoms, with the first three values corresponding to concepts are initially set to 0. The rest of the concepts' values are based on the patients' symptoms, which are given in a hazy manner.

Step 4: With the aid of the weight matrix, the initial vector is employed to achieve the equilibrium state. As illustrated in, a sigmoidal non-linear function is employed as a threshold function (21)

$$f(x) = \frac{1}{1+e^{-\lambda x}} (21)$$

Step5: After a specific number of iterations, the procedure is repeated until an equilibrium state is reached. The most likely identification is the largest value among the values of records.

Step 6: If the readings are high and the difference is less than 10%, FCM alone will not be able to make the correct diagnosis. As a result, the NCM model provides a difference of more than 10% between the leading choice ideas, indicating the most likely diagnosis.

Because the NCM model is limited in its ability to manage indeterminacy in real-world data, the MBA-NCM model was suggested.

- **Basic Principle of Bat Algorithm**

BAs (Bat Algorithms) are a ground-breaking new swarm intelligence optimization approaches that mimic bat's foraging behaviours. They are based on the bat's superior echolocation capabilities. The bats (the primary little bats) produces loud and brief pulse sounds that serves as sonar. When sounds strike an item, the echoes swiftly return to bats' ears, allowing them to receive and locate preys. The biological process of BAs, detailed below [28] simulates foraging activities of bats. Echolocations are used by bats to compute distances where detecting obstacles from preys are difficult. Bats search for preys with particular velocities V_i at positions X_i based on varying wavelengths, loudness A_0, and fixed frequencies f_{min}. Wavelengths of bat pulses are modified according to the distances between preys and bats. As they near preys, frequencies of transmissions $r \in (0, 1)$ also get modified. During searches, loudness change from max value A_0 to min value A_{min}.

BAs were inspired by the fascinating behaviour of micro bat echolocations and incorporate existing methods as well as innovative characteristics. Based on these assumptions, these approaches create random collections of solutions and then use loop searches to discover best options. The local searches are currently being utilised implying local solutions are generated by random flights around optimum solutions, which result in global optimal solutions. When bats are within foraging spaces in d-dimensional spaces at t-1 then their locations of i is X_i^{t-1}, while flight velocities are represented by V_i^{t-1}, and X∗ represents current global optimal positions. Hence, positions and flight velocities of bats i at time t are updated using:

$$f_i = f_{min} + (f_{max} - f_{min})\beta, (22)$$

$$V_i^t = V_i^{t-1} + \left(X_i^{t-1} - X\right) f_i, (23)$$

$$X_i^t = X_i^{t-1} + V_i^t, (24)$$

where the sound waves generated by bats have minimum frequencies of f_{min} and maximum frequencies of $f_{max,}$ while β represents uniformly distributed random numbers in the interval [0, 1]. The frequencies of bat's emitted sound waves get uniformly distributed in the range [f_{min}, f_{max}] during the initial setting processes. Equation (1) is used to get the relevant frequency, and then equations (23) and (24) are used to do the local search (24). The bat walks at random according to the ideal solution, which is generated by the following equation..

$$X_{new} = X_{old} + \varepsilon A\acute{(}t) (25)$$

where ε represents random numbers in the range [1, 1], Xold implies solutions chosen at random from current optimal solutions, and A(t) stands for average loudness of bats and t denotes count of rounds. After determination of bat's loudness A_i and rate r_i, updates can be defined as follows. If bats detect a prey, their pulsed emissions will become less responsive and rates at which they emits pulses raise [29]. The following equations update the bat launch pulse's loudness A_i and rate r_i

$$r_i^{A_i^{t+1}} = \alpha A_i^t, \quad (26)$$

$$yt$$

$$1 + exp, (27)$$

$$r_i^{t+1} = r_i^t$$

where r0 i is the initial rate and A0 i is the initial loudness, which are all randomly chosen. α and y are constants ($0 < \alpha < 1$, $y > 0$).

- **Improved Steady-state Distribution based Bat Algorithm (ISDBA)**

This technique is improved in this work by using feature selection to choose the best characteristics in a random manner. For the selection of relevant features from the dataset, a bat optimization approach is provided. This is one of the optimization techniques that improves the accuracy of final findings by optimising the attributes of the dataset. Microbats utilise echolocation to detect their prey and mates in general. Bats produce a sound pulse in the range of 20Hz to 150Hz to detect obstructions in their path of passage and to determine their destination. On the receipt of the sound pulses, barriers or prey reflect an echo. The bat determines whether the echo is from the prey or the obstacle based on the signal intensities of the received echo. The intensity of signals may also be used to calculate the distance and position of objects. Similarly, vital characteristics are chosen from the dataset based on the kinds of attributes, while unnecessary features are excluded by treating them as roadblocks. The process of finding the best characteristics for diseased rice crops using ISDBA is depicted in Figure 2.

The dataset is regarded the original Bat population in the enhanced bat algorithm, and each data has a specific frequency f_i and velocity v_i. The following equations are used to estimate frequency and velocity, and they are updated after each iteration..

$$f_i = f_{min} + (f_{max} - f_{min})\delta (28)$$

$$v_i^t = v_i^{t-1} + (G_i^t - G_{curren})f_i(29)$$

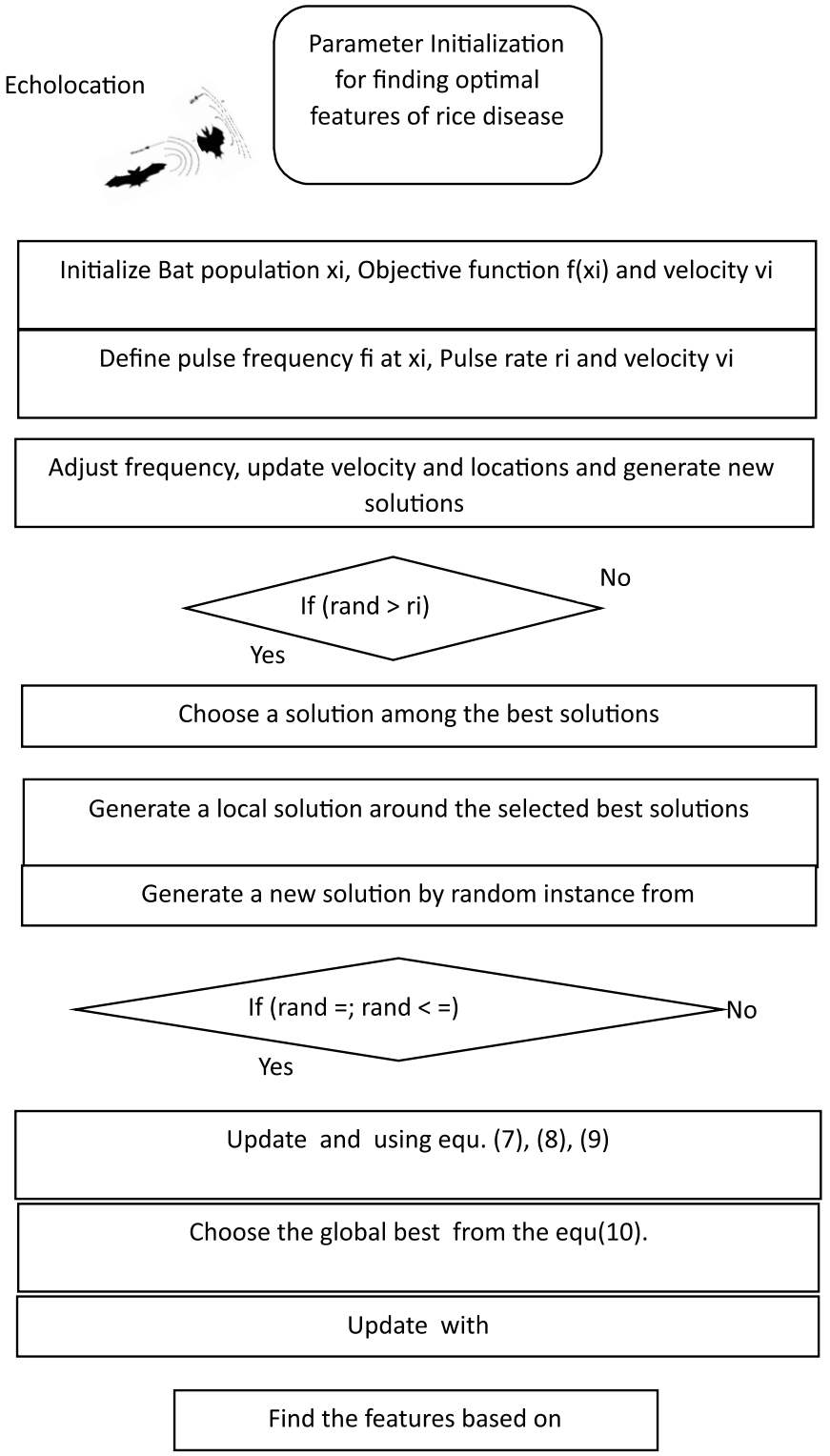
$$\delta = \delta_0 exp(-\mu k)(30)$$

Where, δ_0 and μ are constants and k is the current generation. And $G_{current}$ represents the present global solution and δ denotes the Steady state distribution function that ranges between 0 and 1.

Then, using random walks inside the dataset, local search is used to find the best solutions. The main characteristics from the full dataset, which comprises a vast number of attributes, are extracted using a simple random sample procedure. The characteristics are combined together and sorted using a resampling process based on their similarity. The current solution is compared to the ranking solutions, and the best solutions are arranged first. The following equation is used to update the solutions.:

$$G_{best} = G_{current} + \epsilon S_i(31)$$

where, ϵ is a random number that ranges between -1 to 1 and S_i signifies the similar features.



Figurre .2. Process of selecting the optimal features for diseased rice crops using ISDBA

3.7. Classification using ECMs

Ensemble methods are meta-algorithms that integrate many machine learning approaches into a single prediction model and offer enhanced predictions when compared to single prediction models. As a result, ensemble approaches have won first place in a number of important machine learning contests. Ensemble approaches have had a lot of success in breaking records on difficult datasets. Three machine learning-based classifiers are employed to classify the data in this study. Here the MGNNs , Extreme Gradient Boost Classifier and Logistic Regression (MLR). Finally, output the classification result based on the voting method.

3.7.1. MGNNs

GNNs (Granular neural networks) are ANNs (artificial neural networks) that can handle data that is quantitative or granular in nature. GNNs are based on incremental [Online learning from data streams 30]. It is evident from Fig. 3, learning in GNNs are based on same notions. Initial granules of information i.e. fuzzy sets - are created using original numeric representations. The NNs (neural networks) learning are based on the formed information granules instead of original inputs. Thus, NNs get exposed to relevant data formed where instances are deleted if they don't have any new information.

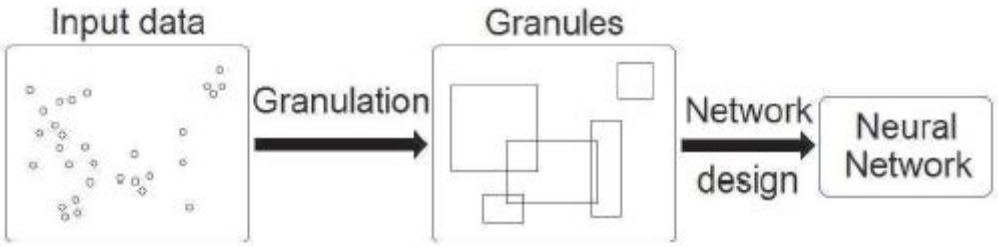


Figure 3. Two-stage design of granular neural networks

GNN analyze data streams using fast incremental one pass learning approaches. They start from zero information on statistical properties of data for classifications and proceed towards granulating feature spaces with fuzzy hyperboxes. Important characteristics of GNNs that broadcast news globally are:

- Deal with labelled and unlabeled examples at the same time;
- Detect drifts and deal with data uncertainties;
- Demonstrate non-linear separation capabilities; and
- Develops life-long learning using constructive bottom-up and destructive top-down mechanisms.

a) GNN Structure and Processing

GNNs learn from data streams $x[h], h = 1,2,..$.Training examples may be accompanied by classlabels $C[h]$.Information granules $\gamma_i$of finite collectionsof granules $\gamma = \{\gamma 1, \ldots, \gamma c\}$in feature spaces $X \subseteq Rn$are associated with class Ck from finite collections of classes$C = \{C1, \ldots, Cm\}$in output spaces $Y \subseteq N$. GNNs link features and output spaces using granules extracted fromdata streams and layers of T-S neurons. NNs have 5-layer structures as shown in Fig. 4. Input layers fan feature vectors $x[h] = (x1, \ldots, xj, \ldots, xn)[h], h = 1, \ldots$ into networks; granular layers consist information granule rule sets $\gamma i \forall i$formed within scopes of feature spaces. Granules are allowed to partially overlap; aggregation layers encompass null neurons $TSni \forall i$ and aggregate membership values to generate values $oi \forall i$representing class compatibilities between examples and granules; decision layers compare compatibility values o_i, and classes $\acute{C}_k$associated with granules $\gamma_i$with highest compatibility values are output; output layers are class label indicators. All layers, except the input layer, evolve as $x[h], h = 1, \ldots$, are inputs.

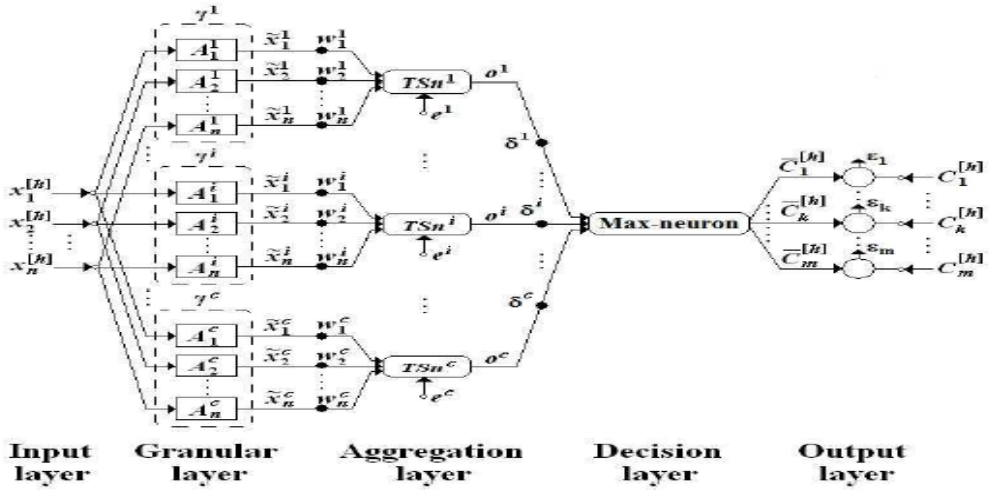


Fig. 4. Structure of the evolving granular neural network for classification

Depending on the application, structural and parametric customisation of the GNN classifier may be accomplished in a variety of methods. When the number of courses, for example, is known ahead of time, it can be automatically adjusted. If memory and processing time are constrained, the number of granules in the model structure may be constrained as well. In new circumstances, the learning algorithm may automatically regulate the number of granules and classes.

SBOAs (symmetric butterfly optimization algorithms) are utilised to give a novel optimization method for GNNs where GNNs use count of modules or sub-granules, trained data proportions, learning rates, goal errors, hidden layers counts, and neurons counts in layers. There are several optimization strategies available today, but it is critical to select one that is suited for the application and allows for better outcomes. As a result of this research, SBOAs were developed for optimising parameters of GNNs.

b) BOAs (Butterfly optimization algorithms)

The butterflies are idealised in terms of a search algorithm to convey the aforementioned concepts: 1. All butterflies are programmed to release a smell that attracts other butterflies. 2. Butterflies fly in any direction or towards odour-producing butterflies. 3.Sensory intensities of butterflies are influenced by objective functions. The three stages of BOAs are: (1) launch, (2) iteration, and (3) completion.

The startup phases are completed in the runs of BOAs, followed by iterative searches, and terminated on finding optimal solutions. In the first step, BOAs determine target functions and their solution spaces with assignment of values to parameters. On selection of variables, initial population of butterflies are generated for optimizations. Since, count of butterflies in simulations of BOAs are constant, a specific amount of memory is set aside for storing their data. In search spaces, butterflies are generated at random, and their scent and fitness values are computed and recorded. On initialization, BOAs move towards iterations where fake butterflies are created.

In iteration phases, BOAs perform a specific number of iterations. All butterflies in solution space are relocated to new locations and their fitness values are determined with each iteration. To begin, the method calculates all of the butterflies' fitness values at different points in the solution space. The butterflies will then produce scent at their respective locations using Eq (32). There are two phases to the technique: global search and local search. During the global search phase, butterflies get closer to the fittest butterfly/solution g, which can be represented by Eq (32)

$$x_i^{t+1} = x_i^t + \left(r^2 \times g \quad - x_i^t\right) \times f_i (32)$$

where x_i^t stands for ith butterfly's solution vector in iteration t. g∗ stands for current best solution amongst all solutions in the current iteration. f_i represents ith butterfly's fragrance while r stands for random numbers in the interval [0, 1]. Local searches are represented as

$$x_i^{t+1} = x_i^t + \left(r^2 \times x_j^t - x_k^t\right) \times f_i (33)$$

where x_j^t stands for jth butterfly, x_k^t represents kth butterfly. If $x_j^t$and x_k^t are from the same swarm and random number r in the range (0,1) Equation (33) get transformed into local random walks. Butterflies look for food and mates both locally and globally. When physical proximity and other factors such as rain, wind, and other factors are considered, the searches for food can account for significant portions of butterfly's mating or food-finding activities. To switch from common global searches to intense local searches in BOAs, switch probabilities p are used.

The traditional BOAs are capable of properly solving the issues. However, it has numerous drawbacks, including early convergence, a proclivity towards local optima, and poor performance. The symmetric distraction factor (SDF) is linked with the BOA algorithm to alleviate the difficulties of BOA, resulting in the S-BOA method..

- **Symmetric Distraction Factor (SDF)**

To overcome these drawbacks, this work presents an modified algorithm. In this approach, a kind of relationship named symmetric distraction factor its neighboring data toward its own data. This symmetric distraction factor depends on two factors; the data intensities or feature attraction $\lambda(0 < \lambda < 1)$, and the spatial position of the neighbors or distance attraction $\xi(0 < \xi < 1)$, which also depends on the neighborhood structure. Considering this symmetric distraction factor defined as below

$$sd^2(x_j, v_i) = \left\|x_j - v_i\right\|^2 \left(1 - \lambda H_{ij} - \xi F_{ij}\right) (34)$$

where H_{ij} represents the feature attraction and F_{ij} represents the distance attraction. The parameters λ and ξ adjust the degree of the two neighborhood attractions.Finally the MGNN model provides the best results.

3.7.2. Extreme Gradient Boost (XGB) classifier

XGBoost (Extreme Gradient Boosting) is a regression and classification machine learning algorithm based on GBDTs (Gradient Boosting Decision Trees) [31]. The centre nodes of the regression tree include attribute test values, while the leaf nodes with scores indicate a choice. Because XGBoost is an additive learning approach with a second-order approximation, the first-order and second-order derivatives of the loss functions with respect to the prediction are required for model fitting.

Second-order approximation of additive tree boosting is derived to provide a clear technique, where m represents the quantity of data and n represents the number of features. The sigmoid function's 'raw prediction' will be represented by zi, and the probabilistic forecast by $\hat{y}_i = \sigma(z_i)$, The sigmoid function is represented by $\sigma(.)$. It's important to remember that the notations differ, and as a result, there's a discrepancy between notations and $\hat{y}_i$ in analysis is denoted as z here. y_i is used to denote the true

label, $\alpha$and γ and are utilised to define the parameters for the two loss functions [32]. The expressions of the gradients/hessians are recorded in a merged format that is independent of the value of y_i, which can make programme implementation easier and aid vectorization in other programmes. The additive learning target that is used in practise is:

$$L^{(t)} = \sum_{i=1}^{n} l\left(y_i, z_i^{(t-1)} + f_t(x_i)\right) + \Omega(f_t) \quad (35)$$

where t denotes the t-th iteration of the training process. Notice that the replacement of the notations has been applied in the equation. Applying second-order Taylor expansion on equation 8, one will get:

$$l\left(y_i, z_i^{(t-1)} + g_i f_t(x_i)\right) + \frac{1}{2} h_i \left(f_t(x_i)\right)^2 + \Omega(f_t) \quad (36)$$

$$L^{(t)} \approx \sum_{i=1}^{n} g$$

$$\left[if\right| \quad \left| t(x_i) + \frac{1}{2} h_i \left(f_t(x_i)\right)^2\right] + \Omega(f_t) \quad (37)$$

$$\propto \sum_{i=1}^{n}$$

The last line comes from the fact that the $l\left(y_i, z_i^{(t-1)}\right)$ term can be removed from the learning objective as it is unrelated to the fitting of the model in the t-th iteration.

Hand-derived derivatives will be required because XGBoost does not support automatic differentiation. In the meantime, the generated expressions could be used in a variety of machine learning applications. Both loss functions have sigmoid activations, and the following fundamental sigmoid property will be used throughout the derivatives:

$$\frac{\partial \hat{y}}{\partial z} = \frac{\partial \sigma(z)}{\partial z} \quad (38)$$

$$\sigma(z)\left(1 - \sigma(z)\right) \quad (39)$$

$$\hat{y}(1 - \hat{y}) \quad (40)$$

And it is a best method for identifying the diseased rice crops in an efficient manner.

3.7.3. LRs

The logistic regression model is used, which is a fairly prevalent model in machine learning and is frequently used in real-world applications such as data mining. For example, in this study, the risk variables for provided data were addressed, and the chance of occurrence was predicted based on the risk factors. The most common use of Logistic regression is used to forecast the chance of each classification event occurring in two-category problems (that is, there are only two forms of output, each representing one category). A logistic regression model is depicted in the picture below:

$$prob(Y = 1) = \frac{e^z}{1 + e^z} \quad (41)$$

Where Y refers to binary dependent variable (Y is equal to 1 if event happens; Y=0 otherwise), e stands for the foundation of natural logarithms and Z means $Z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots \ldots \ldots \ldots \ldots \ldots + \beta_p X_p$ with constant β_0 , coefficents β_j and predictors X_j , for p predictors (j=1,2,3,.....,p).

4. Results and Discussion

The following experiments were performed, and their results are reported, to evaluate the effectiveness of the proposed incremental classifier on sick rice crop data for predicting different rice diseases, as well as on particular benchmark data. The diseased rice crops dataset was used in this study for the prediction of several rice illnesses. Because fresh diseased rice crops data is supplied to the old data on a daily basis, an effective incremental classification approach is required to develop updated classification rules in this dynamic context. The proposed classification model was applied to a diseased rice crops dataset to provide revised classification rules for the prediction of various rice illnesses using the ECM approach..

Poisoning from Rice The fungus PyriculariaGrisea causes rice blast, the disease Sarocladiumoryzae causes sheath rot, the fungus Bipolarisoryza causes leaf brown spot, and the bacterium Xanthomonasoryzae causes bacterial blight. As training data and first classification rules are established, 60% of the dataset learns a classification model. The remaining data is then utilised as incremental data, with the remaining 20% used as test data to assess the system's performance in a dynamic setting.

4.1. Performance of the classifier

To assess how well the proposed ECM technique worked, it was tested on five benchmark datasets from the UCI repository (UCI, 2010). These datasets, like the damaged rice crops dataset, are divided into training, new, and test sets, and various classifiers are trained and evaluated. Table 3 compares the proposed ECM technique's classification accuracies to those of many other classifiers previously used to classify damaged rice crops. All existing classifiers are run in a static environment where the full dataset is analysed at once and accuracy is assessed using a 10-fold cross validation technique. The recommended ECM-based classifier is

only usable in dynamic settings. Aside from classification accuracy, several statistical measurements are performed to evaluate the classifier as shown in equations (42)–(45), and the average is calculated.

Precision is defined as the ratio of correctly found positive observations to all of the expected positive observations.

$$\text{Precision} = TP/TP+FP \quad (42)$$

Sensitivity is defined the ratio of correctly identified positive observations to the over-all observations in real class – yes.

$$\text{Recall} = TP/TP+FN \quad (43)$$

F1 score is defined as the weighted average of Precision as well as Recall. As a result, it takes false positives and false negatives.

$$\text{F1 Score} = 2*(\text{Recall} * \text{Precision}) / (\text{Recall} + \text{Precision}) \quad (44)$$

Accuracy is calculated in terms of positives and negatives as follows:

$$\text{Accuracy} = (TP+FP)/(TP+TN+FP+FN) \quad (45)$$

Where TP, FP, TN, FN are defined as True Positive, False Positive, True Negative, False Negative respectively. These parameter values are calculated for all the standard and proposed classifiers for all five benchmark datasets (UCI, 2010) and a diseased rice crops dataset.

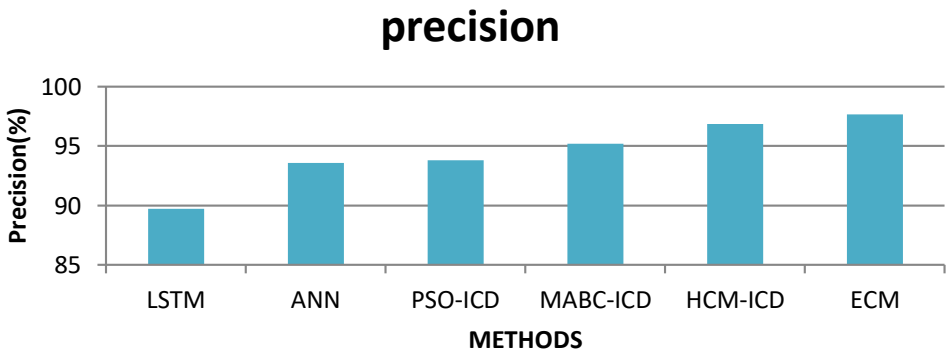


Figure .5. Precision comparison results of the proposed ECM algorithm based Incremental classification

Figure 5 depicts the performance of the suggested ECM's accuracy comparison findings. As a consequence, the findings show that employing a Hybrid Neutrosophic Cognitive Map to select characteristics can be helpful in predicting rice illness categorization. As a result, the proposed HNCM contains a large number of valuable characteristics that have no impact on the performance of the jointly learned linear transformation. It is a desirable trait since it eliminates the need to painstakingly modify the regularisation parameter in the classifier. For tackling the unsupervised classification problem, the suggested HNCM offers a highly successful method..

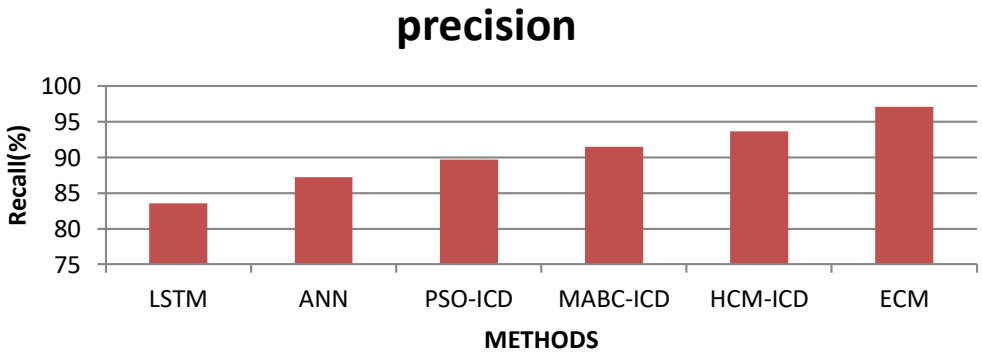


Figure .6. Recall comparison results of the proposed ECM algorithm based Incremental classification

The performance results of the suggested ECM method are shown in Figure 6. Thus, the suggested approach provides high recall results, whereas the current methodology provides lower recall results, such as HCM-ICD method metric of 93.64 percent and MABC-ICD method metric of 91.74 percent, respectively. The PSO-ICD method metric has an accuracy of 89.68 percent, the ANN method metric has an accuracy of 87.25 percent, and the LSTM method metric has an accuracy of 83.54 percent..

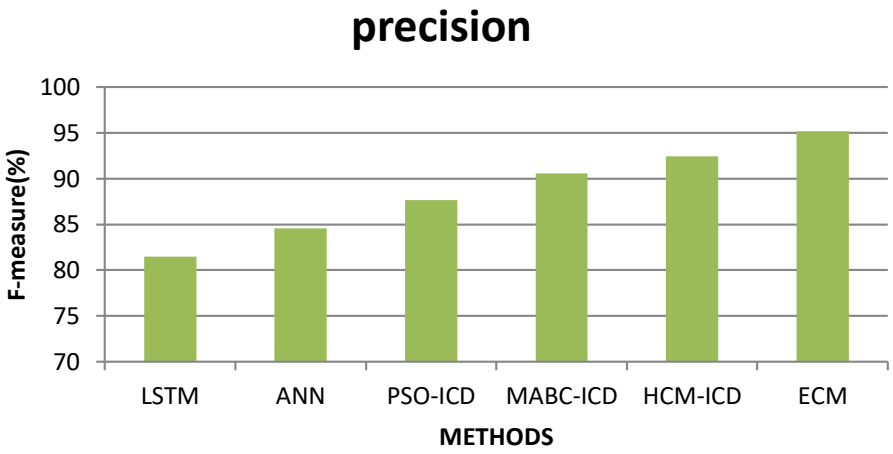


Figure .7. F-measure comparison results of the proposed ECM algorithm based Incremental classification

In comparison to the HCM-ICD, MABC-ICD, PSO-ICD, ANN, and LSTM, the suggested incremental learning paired with the ECM classifier produces a remarkable performance in terms of illness prediction rate, as shown in figure 7. The quantitative analysis' F-measure results are consistent with the qualitative analysis' incremental learning outcomes. The proposed ECM is compared with other state-of-the-arts classification algorithms in terms of their accuracies for diseased rice crops dataset.

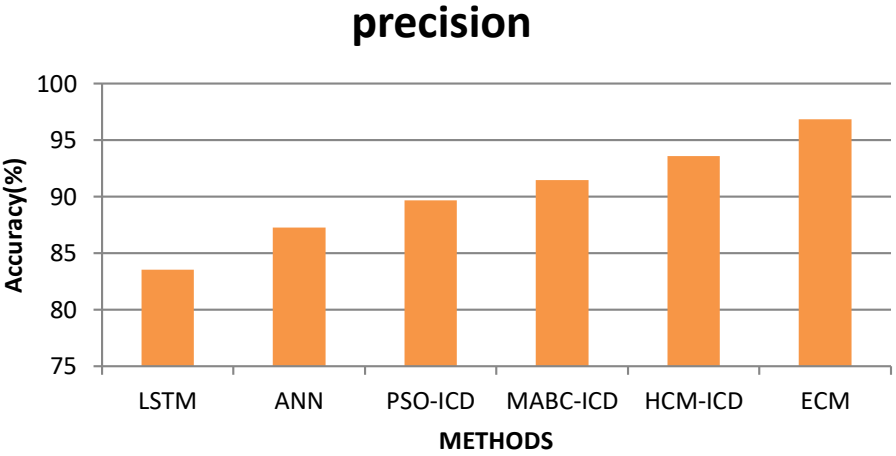


Figure .8. Accuracy comparison results of the proposed ECM algorithm based Incremental classification

Figure.8 indicates that the proposed ECM-based incremental classifiers outperform the current classifier in terms of accuracy. Similarly, all of the aforementioned classifiers are tested in a static context and perform poorly when compared to the ECM classifier, demonstrating the method's efficiency in a dynamic setting for diseased rice crops classification. As a result, when compared to incremental classifiers constructed on previously created models and fresh data that is similar to the entire dataset, the accuracy of the classifiers will be higher..

5. Conclusion

It is critical to detect any illness early and perform the necessary treatment to the damaged plants in order to ensure the rice plants' healthy and appropriate growth. This paper provides a diseased rice crops detection system that employs Hybrid Feature Selection and Ensemble Machine Learning techniques for accurate disease diagnosis. For data normalisation, SMOTE-based pre-processes are recommended first. Then, for efficient segmentation, the MFWFC based segmentation is developed. Then, to improve the classifier's performance, the EICAs are created. The Hybrid Neutrosophic Cognitive Maps are used to choose the features (HNCM). Finally, the ECMs were created in order to improve prediction accuracy. The feature extraction procedure based on a hybrid model extracts features more successfully than existing feature extraction methods. According to the confusion matrixes study, the Ensemble Classification Model, This combination of these three sub-models resulted in proper judgement of confusable diseases. The Ensemble Model has a 91% accuracy rate, indicating that it has sufficient generalisation power to be employed in the identification of damaged rice crops in the field. This demonstrates how the Ensemble Model addresses the issue of a single model misclassifying some disorders by combining the advantages of each model. Because software designs includes both servers and clients, the proposed Smartphone app provides high accuracy, ease of use, simplicity, and a low-cost method of detecting rice infections. The Ensemble Model has a lot of variables that can affect the speed of identification. The focus of future research will be on network pruning to reduce the number of parameters.

REFERENCES

1. Al-Amin, M., Karim, D. Z., & Bushra, T. A. (2019, December). Prediction of rice disease from leaves using deep convolution neural network towards a digital agricultural system. In *2019 22nd International Conference on Computer and Information Technology (ICCIT)* (pp. 1-5). IEEE.
2. Bera, T., Das, A., Sil, J., & Das, A. K. (2019). A survey on rice plant disease identification using image processing and data mining techniques. In *Emerging Technologies in Data Mining and Information Security* (pp. 365-376). Springer, Singapore.
3. Prabhakaran, A., Lekshmi, K. C., & Janarthanan, G. (2021). Advancement of data mining methods for improvement of agricultural methods and productivity. In *Artificial Intelligence in Data Mining* (pp. 199-221). Academic Press.
4. Alfred, R., Obit, J. H., Yee, C. C. P., Haviluddin, H., & Lim, Y. (2021). Towards Paddy Rice Smart Farming: A Review on Big Data, Machine Learning and Rice Production Tasks. *IEEE Access*.
5. Dey, U. K., Masud, A. H., & Uddin, M. N. (2017, February). Rice yield prediction model using data mining. In *2017 International Conference on Electrical, Computer and Communication Engineering (ECCE)* (pp. 321-326). IEEE.
6. Gandhi, N., & Armstrong, L. J. (2016, December). A review of the application of data mining techniques for decision making in agriculture. In *2016 2nd International Conference on Contemporary Computing and Informatics (IC3I)* (pp. 1-6). IEEE.
7. Kumar, S. R., & Kumar, K. R. (2015). A Study on Paddy Crops Disease Prediction Using Data Mining Techniques. *crops*, *1*, 2.
8. Gandhi, N., Armstrong, L. J., & Nandawadekar, M. (2017, April). Application of data mining techniques for predicting rice crop yield in semi-arid climatic zone of India. In *2017 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR)* (pp. 116-120). IEEE.
9. Nithya, A., & Sundaram, V. (2011). Classification rules for Indian Rice diseases. *International Journal of Computer Science Issues (IJCSI)*, *8*(1), 444.
10. Daniya, T., & Vigneshwari, S. (2019). A review on machine learning techniques for rice plant disease detection in agricultural research. *system*, *28*(13), 49-62.
11. Kodama, T., & Hata, Y. (2018, October). Development of classification system of rice disease using artificial intelligence. In *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 3699-3702). IEEE.
12. Hsieh, J. Y., Huang, W., Yang, H. T., Lin, C. C., Fan, Y. C., & Chen, H. (2019). Building the rice blast disease prediction model based on machine learning and neural networks. *EasyChair: Manchester, UK*.
13. Das, A., Mallick, C., & Dutta, S. (2020). Deep learning-based automated feature engineering for rice leaf disease prediction. In *Computational Intelligence in Pattern Recognition* (pp. 133-141). Springer, Singapore.
14. Bashir, K., Rehman, M., & Bari, M. (2019). Detection and classification of rice diseases: An automated approach using textural features. *Mehran University Research Journal of Engineering & Technology*, *38*(1), 239-250.
15. Pinki, F. T., Khatun, N., & Islam, S. M. (2017, December). Content based paddy leaf disease recognition and remedy prediction using support vector machine. In *2017 20th International Conference of Computer and Information Technology (ICCIT)* (pp. 1-5). IEEE.
16. Sharma, R., Das, S., Gourisaria, M. K., Rautaray, S. S., & Pandey, M. (2020). A model for prediction of paddy crop disease using CNN. In *Progress in Computing, Analytics and Networking* (pp. 533-543). Springer, Singapore.
17. Sabitha, R., & Kiruthiga, G. (2021). Plant Disease Recognition and Clustering Using Fuzzy Algorithm on Data Mining. *ICTACT Journal on Soft Computing*, *11*(4), 2429-2432.
18. Karthick, S., Yuvaraj, N., Rajakumari, P. A., & Raja, R. A. (2021). Ensemble Similarity Clustering Frame work for Categorical Dataset Clustering Using Swarm Intelligence. In *Intelligent Computing and Applications* (pp. 549-557). Springer, Singapore.
19. Das, S., Sengupta, S., & Das, P. (2022). Feature Selection Using Louvain Clustering Algorithm for Rice Disease Prediction. In *Computational Intelligence in Pattern Recognition* (pp. 99-108). Springer, Singapore.
20. Krishnamoorthy, N., Prasad, L. N., Kumar, C. P., Subedi, B., Abraha, H. B., & Sathishkumar, V. E. (2021). Rice leaf diseases prediction using deep neural networks with transfer learning. *Environmental Research*, *198*, 111275.
21. Sengupta, S., & Das, A. K. (2017). Particle Swarm Optimization based incremental classifier design for rice disease prediction. *Computers and Electronics in Agriculture*, *140*, 443-451.
22. Deng, R., Tao, M., Xing, H., Yang, X., Liu, C., Liao, K., & Qi, L. (2021). Automatic diagnosis of rice diseases using deep learning. *Frontiers in Plant Science*, 1691.
23. Sengupta, S., & Das, A. K. (2017). Particle Swarm Optimization based incremental classifier design for rice disease prediction. *Computers and Electronics in Agriculture*, *140*, 443-451.
24. Patro, S., & Sahu, K. K. (2015). Normalization: A preprocessing stage. *arXiv preprint arXiv:1503.06462*.
25. Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, *16*, 321-357.
26. Hyvärinen, A., Hurri, J., & Hoyer, P. O. (2009). Independent component analysis. In *Natural Image Statistics* (pp. 151-175). Springer, London.
27. Antepara, E. J. H., Gamboa, J. E. A., Mndez, M. R. C., & Gonzlez, M. E. P. (2017). Competencies Interdepencies Analysis based on Neutrosophic Cognitive Mapping. *Neutrosophic Sets and Systems*, *16*, 89-92.
28. Yang, X. S. (2011). Bat algorithm for multi-objective optimisation. *International Journal of Bio-Inspired Computation*, *3*(5), 267-274.
29. Gandomi, A. H., Yang, X. S., Alavi, A. H., & Talatahari, S. (2013). Bat algorithm for constrained optimization tasks. *Neural Computing and Applications*, *22*(6), 1239-1255.

30. Song, M., & Pedrycz, W. (2013). Granular neural networks: concepts and development schemes. *IEEE transactions on neural networks and learning systems*, *24*(4), 542-553.
31. Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., & Chen, K. (2015). Xgboost: extreme gradient boosting. *R package version 0.4-2*, *1*(4), 1-4.
32. Babajide Mustapha, I., & Saeed, F. (2016). Bioactive molecule prediction using extreme gradient boosting. *Molecules*, *21*(8), 983.