Ensemble Modulation Pattern based Paddy Crop Assist for Atmospheric Data

SampathKumar.S.[1+], B.N.ManjunathaReddy[2++] and M.Nataraju[3+++]
sampathkumars@ksit.edu.in, manjunatha_reddy@gat.ac.in, nati_setty1975@rediffmail.com
Department of ECE, K.S.I.T., Global Academy of Technology, VTU,
GPT, Srirangapatna, Karnataka,560098,India.

Summary

Classification and analysis are improved factors for the real-time automation system. In the field of agriculture, the cultivation of different paddy crop depends on the atmosphere and the soil nature. We need to analyze the moisture level in the area to predict the type of paddy that can be cultivated. For this process, Ensemble Modulation Pattern system and Block Probability Neural Network based classification models are used to analyze the moisture and temperature of land area. The dataset consists of the collections of moisture and temperature at various data samples for a land. The Ensemble Modulation Pattern based feature analysis method, the extract of the moisture and temperature in various day patterns are analyzed and framed as the pattern for given dataset. Then from that, an improved neural network architecture based on the block probability analysis are used to classify the data pattern to predict the class of paddy crop according to the features of dataset. From that classification result, the measurement of data represents the type of paddy according to the weather condition and other features. This type of classification model assists where to plant the crop and also prevents the damage to crop due to the excess of water or excess of temperature. The result analysis presents the comparison result of proposed work with the other state-of-art methods of data classification.

Keywords:
Paddy crop classification; Ensemble Modulation Pattern; Block Probability Neural Network; Cultivation

1. Introduction

In the agriculture field, the prediction of humidity with respect to temperature of land guides the selection of crop to cultivate and the irrigation [1]. For that there are several forecasting methods are available based on the history of humidity data for each instant of time samples. Based on the linearity of each data samples, the humidity level was forecasted and according to that the cultivation of crops or the irrigation chart were followed. Based on linearity of data samples, the forecasting result is predicted with the increase in error rate. This needs to improve the prediction rate by implementing an advanced model of feature prediction and classification system [2] to analyze the history data for that area and forecast the relevant parameter values.

Manuscript received September 5, 2022
Manuscript revised September 20, 2022
https://doi.org/10.22937/IJCSNS.2022.22.9.53

To enhance the prediction model and the feature analysis of texture pattern from the input properties of temperature and the humidity data samples are predicted by optimization algorithm [3] to select the best attributes in the paddy data classification process, the temperature and the humidity data samples are used for estimating the type of rice that can cultivate in that atmosphere and to predict the growth level of it. Entire and uniform grain sterility in rice over large areas has become a serious problem in the warm and humid lowland region. In view of this, the influence of high temperature at low and very high relative humidity (RH) levels and normal temperature at very low and normal RH levels on the spikelet surface temperature and grain sterility in rice at heading were studied under controlled environments. Entire grain sterility in rice was induced by high temperature (35°C day/30°C night) when coupled with high RH (85 - 90%) at heading, Reduction in RH by 30%

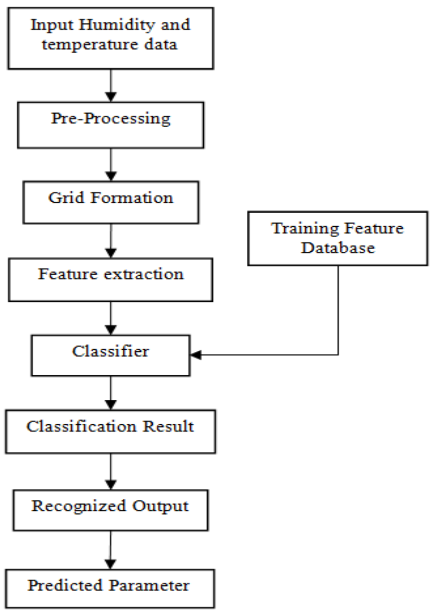


Fig.1 Basic Architecture Diagram of data analysis

at high temperature resulted in decrease in grain sterility but made no significant increase in the completely filled grain percentage which was negligibly low at both very high and low RH levels.

In most commonly, the forecasting of the atmospheric condition [4] is to train more number of data samples to get the best classification performance and also to improve the accuracy of prediction. This form of training part increases the time complexity and the space complexity of the classification model. This can be overcome by optimal feature selection with the texture pattern based data classification approach.

By considering that, the list of objectives that are focused in this paper work are

- To implement a novel Ensemble Modulation Pattern based texture pattern extraction algorithm for representing the feature of input data samples.
- To enhance the classification performance by arranging the optimal feature attributes that are relevant to the texture data.
- To validate the humidity and temperature kind of feature attributes for predicting and forecasting the range of paddy cultivation based on the atmospheric condition.
- To implement the texture classification method by using Block Probability Neural Network classifier.
- To validate the performance of proposed classification model based on the statistical parameters and the comparison result.

The overall description of the paper work with algorithm statements and the comparative study can be organized as in the following subsections. In that, the related works for the texture classification are surveyed and reviewed in the section II. Section III explains about the algorithm description and the steps involved in proposed Block Probability Neural Network classifier-Ensemble Modulation Pattern techniques for paddy data prediction. Section 4validates the performance of proposed work by estimating the statistical parameters and presented as the comparison of classification results with the other traditional classification models. The justification and the future work was concluded in the section 5.

2. Related Work

The detailed review of existing methods in the paddy prediction and the data forecasting are presented in this section. These are all mainly focused on the process of feature extraction, optimization and the classification of data samples that are followed in the flow of work.

In this, the paddy rice pattern prediction was reviewed with the properties of soil level and the irrigation condition based on atmospheric changes. By considering that, [5] presented the study of Azolla compost method to improve the grain yield by mitigating the water deficit stress for the paddy crop. This sub-divides the soil texture and irrigate the water with respect to it. Similarly, based on the atmospheric weather condition, the water supply was predicted and the frequency of water requirement according to the climate change was estimated in [6]. This was tested in the land area of Bangladesh to validate the performance of paddy cultivation. Based on the time series and the data samples of the paddy time chart, the planting date of paddy was estimated in [7]. In that, the MODIS data samples are used for the analysis of production in the basin-scale rice planting model. This also refers the climate change and the weather prediction based on history of dataset. The irrigation schedule was also predicted in [8]. In this work, the short-term weather forecasting data modeling was used to predict the schedule time of irrigation. This was functioning based on the historical data samples of the weather and the soil moister parameters.

In addition to the properties of soil moisture, some other methods predicting the satellite data to estimate the moisture level in that region of interest. With this information, [9] proposed a soil moisture retrieval models based on the prediction of C-band SAR satellite data. This performs a study of different classification model to estimate the soil moisture from satellite data and compare the classification performance based on the statistical features. To validate the effect of rice straw-and rapeseed residue-derived bio-chars, [10] estimates the level that affects the geochemical fractions and other related parameters in the contaminated soil. These are all predicted for the different moisture conditions in the soil based on the parameters of atmospheric state. Similar to that, [11] proposed a prediction model of pH level in the paddy soil based on the conditions of flooding and the drainage of moisture content. This predicts the state of soil that are increasing the range of pH>6.5 to maximum value of 7 after the flooding effect. This estimation helps to improve the paddy cultivation in the appropriate soil. Which this enhancement, [12] proposed the prediction model of water logging in the paddy field to provide the information about irrigation level. Based on the water level, this estimates the amount of water supply and the other related features for better growing the paddy in the field. From these enhancement of the prediction level, [13] proposed a Convolution Neural Network feature estimation based smart farming system. In this, the weed crop was monitored and recognizing the level of growth and based on the graphical representation of convolution network that estimate the state of paddy growth and its conditions.

From the prediction process of soil parameters and the forecasting data, several other methods are improving the cultivation range and the other decision plan. From this, [14] proposed a decision support of estimating the paddy quality based on fuzzy logic. In this, the rice that are stored in the food go-down are classified according to the quality and validate the food range for public distribution. The fuzzy interference system determines the rice quality and segments each division separately. In [15], the paper work proposed the spatial variation estimation in the field of paddy cultivation land. This predicts the pH level of the rice grains during grain fillings. This also estimates the status of soil redox to predict what type of the rice paddy can grow in that soil. To reduce the carbon footprint and to improve the benefit of ecosystem, [16] makes the conversion of double-season rice to ratoon paddy fields. Due to the Ratoon rice (RR) system, this will increase the high annual yield and also required the low cost of fertilization and the labor cost. This was analyzed with different data points in the overall paper with the analytical report of crop cultivation.

To preserve the environment, several methods are focused to estimate the pH level and the other harmful gas emission effect. By considering that, [17] proposed the conversion model of winter flooded paddy planting to rice-wheat rotation. This type of planting the rice and wheat changes the methane emission level due to the reason of high emission in rice planting seasons. Other state of prediction for identifying the plant disease based on the soil nature and the weather conditions are focused in [18]. The disease that are affecting the paddy crop are predicted based on the weather changes and provide the report of paddy level to make the right decision for protecting the crop. In [19], author proposed an intelligent irrigation system to grow the crops based on temperature and soil moisture level that are predicted from the sensor features. From these, [20] proposed a rice paddy distribution prediction based on the remote sensing data. This was estimated and predicted by using the coupling deep learning with phenological characteristics. In [21] author presented a study of different paddy disease classification model to protect the paddy crop.

From these discussion, the texture based soil range prediction are focused in the proposed model based on the soil moisture, temperature of atmosphere and other related features and parameters. This was classified and predicted the type of paddy that can grow in that soil conditions are classified by using the neural network technique and validate the performance of proposed model.

3.Proposed Methodology

The proposed model of temperature and moisture features based on the atmospheric parameters are explained in this section. The major aim focused on this work is to improve the recognition accuracy of forecasting based on the feature representation of the data. For this purpose, a novel technique such as Block Probability Neural Network, and Ensemble Modulation Pattern system techniques are proposed. The architecture of the proposed model of feature classification in the Fig 2 was processed in two functionalities such as,

- Data Preprocessing,
- Pattern Validation and Classification

Initially, the given testing temperature data is preprocessed for normalizing the data with better clustering and optimal selection of features, which is performed by the use of Block Probability Neural Network technique. Then, the blocks of the preprocessed data are segregated for extracting the patterns of the data, where Ensemble Modulation Pattern is applied for extraction. It efficiently extracts the geometrical features for increasing the overall accuracy of classification. Finally, the classifier is employed to classify whether the position ID is satisfied with the relevancy ratio or not.

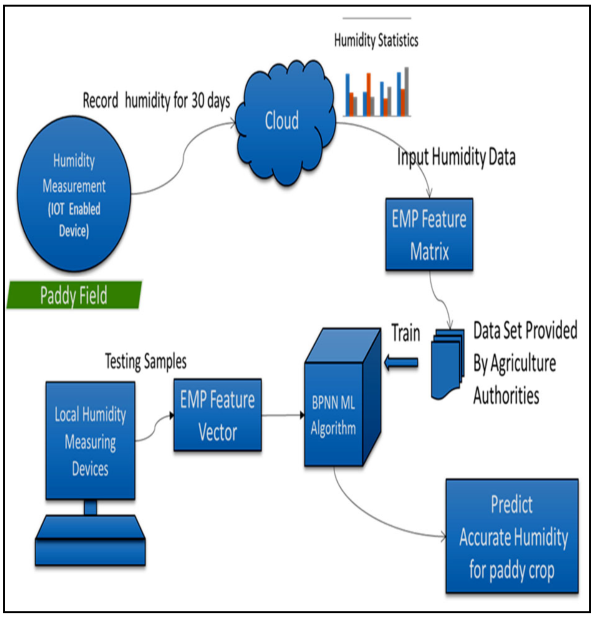


Fig. 2 Flow of the proposed system

3-1-Data Preprocessing

Data preprocessing is a process of preparing the raw data and making it suitable for a machine learning model. It is the first and crucial step while creating a machine learning model. When creating a machine learning project, it is not always a case that we come across the clean and formatted data. And while doing any operation with data, it is mandatory to clean it and put in a formatted way. So for this, we use data preprocessing task.A real-world data generally contains noises, missing values, and maybe in an unusable format which cannot be directly used for machine learning models. Data preprocessing is required tasks for cleaning the data and making it suitable for a machine learning model which also increases the accuracy and efficiency of a machine learning model.

It involves below steps:

- Getting the dataset.
- Importing libraries.
- Importing datasets.
- Finding Missing Data.
- Encoding Categorical Data.
- Splitting dataset into training and test set.
- Feature scaling

Typically, the data preprocessing is one of the initial and important stage in any data processing applications. Because, it is essential to ensure the increased precision rate of the subsequent steps. Also, it reduces the impacts of artifacts that could affect the accuracy of classification. For this purpose, an efficient preprocessing method to form cluster of data or optimal reduction of feature set. It filters the loss of key points and enhances the temperature flow details for obtaining a clear texture patterns for further processing. Moreover, it integrates the data by applying normalization for eliminating an irrelevant and unwanted information. In this model, the loss can be represented as E_{xy} based on the following equation:

$$E_{xy} = \begin{cases} C_{ij}, & if\left(mean\left(T_{ij}\right) > I_{xy}\right) \\ 0, & Otherwise \end{cases} \quad (1)$$

Where, I_{xy} represented the data key points for all x and y.

$x = \{1,2, \ldots M\}$; Where, M is the row size of data.

$y = \{1,2, \ldots N\}$; Where, N is the column size of data. After getting the input data, the sharpening is performed by using the following equation:

$$I_e(x, y) = I_{in}(x, y) + \lambda H(x, y) \quad (2)$$

Where, λ defines the tuning filter parameter and $H(x, y)$represents the high pass filter mask for clustering.

After that, the data is separated into cells based on the following equation:

$$T_{ij} = I_e(x-1; x+1, ; y-1; y+1) \quad (3)$$

Then, the average difference value in T_{ij} is computed with respect to the size of filter mask K and center key point of the mask matrix I_c. Here, the index of mask matrix is represented as shown in Fig 3.In which, the matrix size can be presented in both 3×3 and 5×5 matrix formats. The performance of filtering can be enhanced with respect to the minimum the size of mask. This type of filtering reduces the value of loss ratio value by performing the higher key point reconstruction with the reduced number of loss of key points.

i-1, j-1	i, j-1	i+1, j-1
i-1, j	i, j	i+1, j
i-1, j+1	i, j+1	i+1, j+1

(a). 3 × 3 matrix

i-2, j-2	i-1, j-2	i, j-2	i+1, j-2	i+2, j-2
i-2, j-1	i-1, j-1	i, j-1	i+1, j-1	i+2, j-1
i-2, j	i-1, j	i, j	i+1, j	i+2, j
i-2, j+1	i-1, j+1	i, j+1	i+1, j+1	i+2, j+1
i-2, j+2	i-1, j+2	i, j+2	i+1, j+2	i+2, j+2

(b). 5 × 5 matrix

Fig.3 Indexing of mask matrix for data pre-processing

At last, the pre-processed cluster output of data is taken as $I_f(x, y)$, which is further used for pattern extraction. The working procedure of the proposed Block Probability Neural Network technique is illustrated as follows:

3-2-Pattern Extraction

After preprocessing, the filtered data and are processed by the Ensemble Modulation Pattern system technique. It is also one of the essential stage for extracting the most relevant information that are used for characterize each class on the data. Here, the pattern extraction is mainly performed for increasing the overall accuracy and efficiency of classification. In this algorithm, the filtered data I_f is taken as the input for pattern extraction, in which the zero padding is initialized with size of 2 rows and columns with respect to the border of data pattern. Then, the window sizeI_W can be represented as the mask of 5×5 for extracting the patterns from the input. These data cells can be analyzed in five different angles of the data attributes that are can be indexed as$\{+90^0, +45^0, 0^0, -45^0, -90^0\}$.

Algorithm I – EMP algorithm

Input: Input DataT_D
Output: Features of attributes $F_D(s)$.
Fori = 1 to M //Loop run for 'M' number of iteration.
Initialize attributes 'y'and the weight value 'ω_i'
 Calculate Potential of the attributes P_t^n
 Estimate the likelihood of the attributes by
$$L_{1:i}^m = L_{1:i-1}^m \times L_i^m$$
 Update weight value, $\omega_i(n+1)$
 Update Attributes, $y(n+1)$
 Find maximum likelihood, $m_i^* = \max(L_{1:i}^m)$
 Find maximum relevance value, $\omega_i^*(n)$
If$(m_i^* > m_{i-1}^*)$, then
  Update weight value of attributes and get best relevance value to form feature set.
If $(L_{1:i}^m) > 0$, then
$$s_v = \{s_{v-1}, i\}$$
  End if
Else
  Continue for loop 'i'.
 End If
$$F_D(s) = T_D(s_v)$$
End 'i' Loop

$I_{\alpha_L}(x,y) = \sum_{x=-N_1}^{N1} \sum_{y=-N_2}^{N_2} |I_W(x,y)| \times f_1(\alpha_L, \alpha_U, r)$
(4)

Where, $f_1(\alpha_L, \alpha_U, r) = \begin{cases} 1 & \text{if } \alpha_L \le \alpha_U < r \\ 0 & \text{else} \end{cases}$

$\alpha_L = \{+90^0, +45^0, 0^0, -45^0, -90^0\}, \alpha_U = \alpha_L - 45^0$ // 'r' is the size of mask matrix. According to these index points, the neighboring features of the data samples are can be calculated from the equation (5). This can be represented as in the term of α_k.

$\alpha_k = \{I_M((i-1)\text{to}(i+1), (j-1)\text{to}\,(j+1))\}$ (5)

After that, the set of nearest neighborhood key points are predicted from this boundary key point collections that is for the average value of neighboring features μ_k, which is estimated as

$$\mu_k = \frac{1}{L}\sum_{a=1}^{L} \frac{|\alpha_k(a) - I_M(i,j)|}{I_M(i,j)} \tag{6}$$

Similarly the average difference 'μ_c' between the center value of the mask and the boundaries of it can be calculated as in the equation (7).

$$\mu_c = \frac{1}{L}\sum_{a=1}^{L} \frac{|\alpha_k(a) - I_c|}{I_c} \tag{7}$$

For each iteration, the mean values of μ_k and μ_c are computed based on its sign difference, which is used to extract the binary stream of the separated mask as shown in below:

$$S = \begin{cases} 1, & \text{if}(\mu_k > \mu_c) \\ 0, & \text{Otherwise} \end{cases} \tag{8}$$

After that, the corresponding decimal value of B is computed from the binary streams, which is represented as follows:

$$B = B + (2^{k-1} \times S) \tag{9}$$

Consequently, the maximum key point progression for each key point is estimated as γ_k shown in below:

$$\gamma_k(x,y) = \max_{\alpha_L}\left(I_{\alpha_L}(x,y)\right) \tag{10}$$

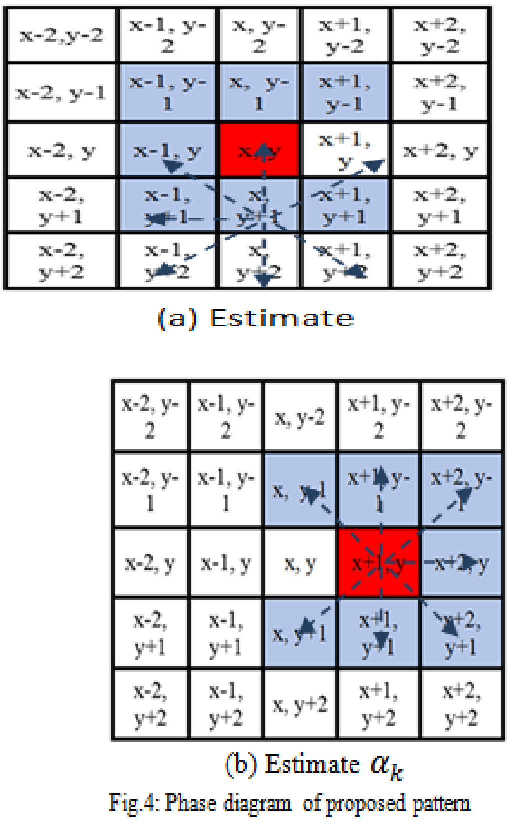


Fig.4: Phase diagram of proposed pattern

Then, the binary code mapping is performed for obtaining the pattern vectors from the input, which are illustrated as follows:

$$I_p(x-2, y-2) = B \oplus \sum_{i=0}^{P} 2^i \times f_2\left(I_{Ref}(s,t), I_{\gamma_1}(s,t), I_{\gamma_2}(s,t)\right) \quad (11)$$

Where, $I_{Ref}(s,t) = I_W(x+t, y+t), \quad \forall t = (-1:1)$

$$f_2(p,q,r) = \begin{cases} 1, & \text{if}(p < r \& q > O \& q > r) \\ 0, & \text{else} \end{cases} \quad (12)$$

The detailed working procedure of the proposed Ensemble Modulation Pattern based pattern extraction is illustrated in Algorithm II.

Fig 4 depicts the phasor representation of Ensemble Modulation Pattern technique, in which the identification of neighborhood key points with respect to various projection angle differences was illustrated. In this, the highlighted red marked region of the metric was used as the reference point to calculate the value of α. For this, the center value of the matrix can be represented asα_c that is estimated and the neighborhood key points for the coordinate position of (x, y) can be represented for computing the value of$'\alpha_k'$. Then, the projection angles of the arrow indicate the estimation of magnitude of current matrix. In this case, the center key point$\alpha_k$is not considered and the boundary key point is not considered for 'α_c'. After extracting the patterns, the geometrical features are also extracted from the filtered data I_f. In this technique, the orientation of the data matrix θ is calculated and the data matrix can be rotated with respect to the updated angle for getting an accurate orientation, which are calculated as follows:

$$\theta = \frac{1}{2}\tan^{-1}\left(\frac{2\mu_{11}}{\mu_{20}-\mu_{02}}\right) \quad (13)$$

Where, μ_{ij} represents the central moments of the data, which is estimated as shown in below:

$$\mu_{ij} = \sum_i \sum_j \left(\left(I_f(x-x_\mu)\right)^i \left(I_f(y-y_\mu)\right)^j\right) \quad (14)$$

Where, the coordinate points 'x' and 'y' for the each data points in the matrix are can be represented asx_μ and y_μ are estimated as follows:

$$x_\mu = \frac{I_f(1,0)}{I_f(0,0)} \quad (15)$$

$$y_\mu = \frac{I_f(0,1)}{I_f(0,0)} \quad (16)$$

Where, $I_f(0,0)$ is the central key point of the data mask, $I_f(0,1)$ and $I_f(1,0)$ are the neighbor key points for the angles of 0^0 and 90^0 respectively. After that, the data matrix can be rotated with respect to the updated for analyzing the correct orientation of the data matrix based on θ, which is shown in below:

$$\theta' = \begin{cases} 90-\theta, & +90 \geq \theta \geq 0 \\ -90-\theta, & -90 \geq \theta < 0 \end{cases} \quad (17)$$

Consequently, the upper peaks and lower peaks in the temperature data is computed based on the previous position and speed of movement. Then, the vertices of the higher peaks and lower peaks are estimated based on the x and y coordinates of the data key points. Based on these, the output geometrical key points F_V are extracted from the filtered data.

Algorithm II –BPNN algorithm

Input: Features matrix of the training set $F_D(s)$

Output: Result from classifier$V(k)$

The initialization of the feature attributes as $F_D(s)$,

$F_D(s) = \{T_{D1}(s), T_{D2}(s), \dots, T_{Dm}(s)\}$ // Initialize the feature properties.

The layers in the neural network are can be defined as the combination of data sequence that are can be represented as $X_D(s) = \begin{bmatrix} F_{D1}(s) \\ F_{D2}(s) \\ \dots \\ F_{Dm}(s) \end{bmatrix}$ // Matrix arrangement for input layer in the Block separation.

The feature blocks are arranged from the matrix values are can be estimated as

$F\left(X_D(s).X_D^*(s)\right)$.

Arrange the kernel function of the classification model, K_m

From the kernel function, the relevancy between the features are calculated as t_n.

$t_n = F^T \omega_n$ // Texture relevancy. 'ω_n' weight value of attributes.

$u_n = F^T \omega_n$ // Texture relevancy.

Calculate the maximum value of matching between the training set and the testing feature attributes as $\widehat{T_s}$

Where, the relevance factor $X_b^{\bar{d}} \in R^{(T-T_p)M}$

Where, 'P' and 'Q^T' – Predicted component.

The predicted label can be representing by

$$V(k) = \frac{d_{ij}}{R_j - R_i}$$

Where, d_{ij} – Distance matrix for 'i' and 'j' of the relevance matrix 'R'.

After extracting the patterns of the historical data of paddy movement, the classification technique is employed for exactly classifying the recognized data. It performs multiple classification processes based on the supervised histogram feature vectors of the data. Here, the texture classification is mainly performed to improve the recognition process by estimating the weight of each neuron in the paddy. This is to predict the network connectivity between the layers of NN and the feature attributes. The relevancy of the temperature and the humidity attributes are can be estimated as the binary value of '0' and '1'. This type of classification can efficiently improve the accuracy rate by using the patterns that are extracted in the earlier stage. Here, the amount of input samples is trained and patterns are classified with increased accuracy rate.

4- Results and Analysis

The results of the proposed model of paddy classification based on the temperature and moisture parameters are validated and compared with the existing methods. The performance of proposed classification model was calculated as the statistical parameters of texture based classification model that are presented as the table result and the graph plot of those parameters. This was implemented and tested in the Python tool in the version of 3.8. For the temperature flow analysis, the position and the coordinates information are selected according to the coverage area and simulated by generating the random position changes and creating the missing of data scenario in different modules. This was validated with the existing methods for the dataset of Portuguese humidity administration database that is referred in paper [22].This contain the temperature information in the Portuguese humidity and in that some data are made as the missing value to predict and forecast the data. Here, the data collection was updating since 2012 for every change in the humidity feature update. For this analysis, there are 2890 collections of humidity sections / instances were arranged to analyze the performance of proposed work comparing with the other state-of-art methods. The performance results are discussed in the following sub-sections.

4-1-Performance Indicators

Performance metrics are a part of every machine learning pipeline. They tell you if you're making progress, and put a number on it. All machine learning models, whether it's linear regression, or a SOTA technique like BERT, need a metric to judge performance.

Every machine learning task can be broken down to either Regression or Classification, just like the performance metrics. There are dozens of metrics for both problems, it is important to know how the model sees your data.

The parameters that are used for the performance metric based analytical process can be calculated by the statistical probability between the truly classified data samples and the number of misclassification result. These are all calculated from the arrangement of confusion matrix that are evaluated by the comparison of classified result with ground-truth of the dataset.

$$\text{Sensitivity, TPR} = \frac{\text{True Positive (TP)}}{\text{Total No.of Positive samples}} \quad (18)$$

$$\text{Specificity, TNR} = \frac{\text{True Negative (TN)}}{\text{Total No.of Negative samples}} \quad (19)$$

$$\text{Jaccard, J} = \frac{TP}{TP+FP+FN} \quad (20)$$

$$\text{Dice Overalap, D} = \frac{2J}{J+1} \quad (21)$$

$$\text{Precision, P} = (1 - FDR) = \frac{TP}{TP+FP} \quad (22)$$

$$\text{Recall, R} = (1 - FNR) \quad (23)$$

$$\text{F1 Score, F\_S} = \frac{2TP}{2TP+FP+FN} \quad (24)$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (25)$$

$$\text{Accuracy, Acc} = \frac{\text{Total correct labels}}{\text{Total No.of Samples}} \quad (26)$$

$$\text{Error } (\%) = (100 - \text{Accuracy}\%) \quad (27)$$

$$\text{Cohen's Kappa} = 1 - \frac{(1-P_o)}{(1-P_e)} \quad (28)$$

Where, P_e – Hypothetical probability and the P_o – Probability of Relative observation,

Fig 5 and Table 1 shows the comparison chart and the table result for the parameters of F1_Score, Mathew's Correlation Coefficient (MCC), Sensitivity, Specificity, and Precision from the reference of [22] and proposed techniques. This represents that the EMP pattern extraction model in the BENN classification model of proposed feature classification method achieved better classification performance than the existing method of TrAdaBoost.

Fig 6 presented the comparison result of accuracy and kappa coefficients for the existing and proposed model of classifiers. From this analysis, it shows that the accuracy of proposed method was increased to 0.983 and kappa coefficient was enhanced to 0.981, when compared to the existing technique. Similarly, the Fig 7 shows comparison result for the parameters of the error rate and FPR. The Table 2 plots the AUC value to represent the texture classification result for proposed paddy prediction model.

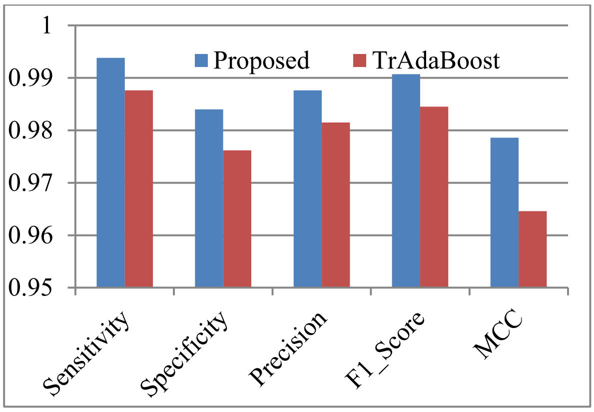


Fig.5 Performance measures

Table 1 Performance evaluation of existing and proposed techniques

Parameters	Proposed	TrAdaBoost
Sensitivity	0.99	0.99
Specificity	0.98	0.97
Precision	0.99	0.98
F1_Score	0.99	0.98
MCC	0.98	0.96
Accuracy	0.98	0.97
Kappa Coefficient	0.98	0.97
Error rate	0.02	0.03
FPR	0.01	0.02

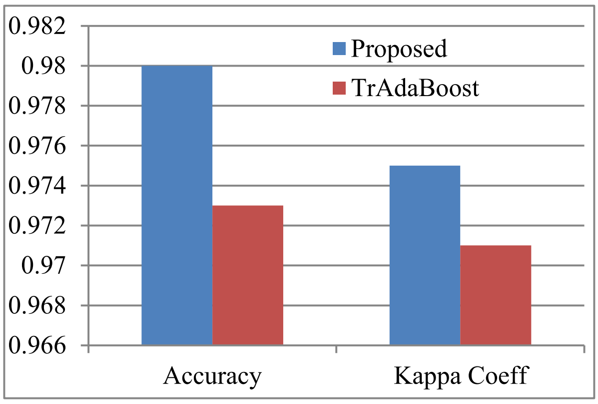


Fig.6 Accuracy and Kappa coefficients

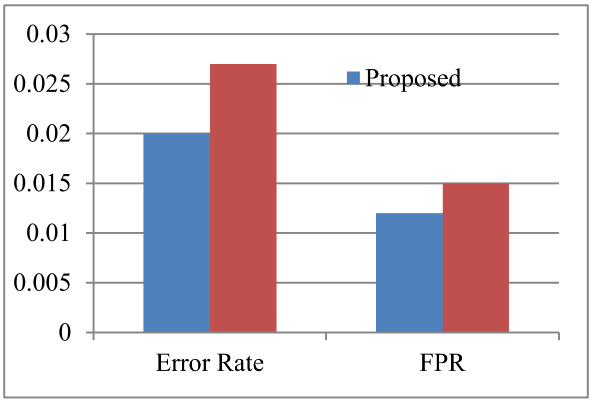


Fig .7 Error rate and FPR

Table 2 AUC analysis

Methods	AUC
MCC	0.601407
CSM	0.572018
MDC	0.865575
Bayes	0.739617
CS	0.937389
TrAdaBoost	0.922684

Proposed	0.965011

4-2-Accuracy Analysis

The Table 3 shows the comparison result of Accuracy from the classification output. In that, the proposed EMP method achieved the better classification result than the other classification model for the parameters of data pattern based feature attributes and the geometrical features of the input temperature and the humidity data of soil. The efficiency of the proposed work can be expressed in the statistical parameters in the range of 0 to 1. This can also have expressed in-terms of percentage by multiplying it into 100 for the ratio value of the probabilistic data analysis. This represents the amount of correctly classified in the EMP method.

4-3- True Positive Rate and False Positive Rate

The true positive rate (TPR) and the false positive rate (FPR) of the referred as the probability of samples that are truly detected as the positive class and the probability of samples that is misclassified as positive class respectively.

Table 3 Accuracy analysis

Methods	Accuracy
MCC	0.85
CSM	0.86
MDC	0.846
Bayes	0.842
CS	0.878
TrAdaBoost	0.953
Proposed	0.98

The comparison chart with respect to the FPR vs TPR was displayed in Fig 8 as the Receiver Operating Curve to represent the efficiency of proposed classification model.

A true positive is an outcome where the model correctly predicts the positive class. Similarly, a true negative is an outcome where model correctly predicts the negative class.

A false positive is and outcome where the model incorrectly predicts the positive class. And a false negative is an outcome where the model incorrectly predicts the negative class.

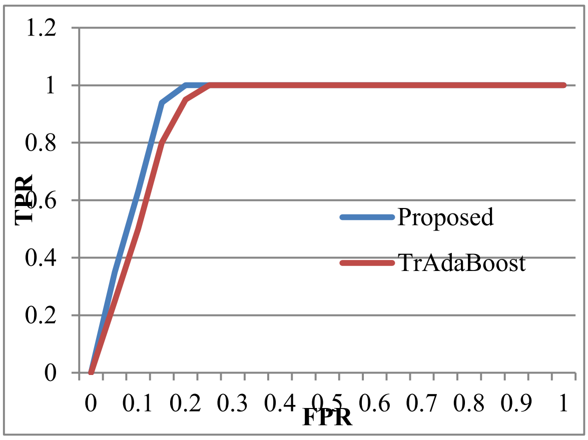


Fig 8. ROC analysis

The overall experimental analysis results depicted that the proposed Ensemble Modulation Pattern –geometric feature extraction based classification technique provides an improved results compared than the other techniques.

5- Conclusion

The proposed work presented a novel pattern extraction based classification method for temperature feature data forecasting and recognition. For this purpose, various data processing techniques are employed in this work at the stages of preprocessing, block separation, pattern extraction, and classification. Initially, the paddy data cluster was implemented to reduce the loss of key points and to integrated with the input data. During this process, the mask matrix is constructed in the form of 3×3 and 5×5, which ensures the reduced loss/ error ratio. After pre-processing the data, the block separation is performed to increase the overall efficiency of recognition. Here, the EMP technique is utilized to extract the most useful patterns by computing the intensity of the center key point with its neighboring key points. This type of feature extract increases the overall accuracy of the data recognition system. At last, the classifier is deployed to classify whether the data is relevant with the prediction or not based on the extracted feature vectors. In this paper, an extensive simulation results have been tested and verified with the performance comparison of predicted result. The traditional classification model was compare with the proposed mechanism, where the results depicted that the combination of Ensemble Modulation Pattern-Block Probability Neural Network technique outperforms the other techniques. Also, it efficiently improves the performance of classification with improved recognition rate, accuracy, and reduced error value.

The future work of this proposed model was can be improved by proposing the optimal classification model with the parameters of temperature and the other soil pattern based texture feature based data forecasting system.

References

[1] Abid, Abbas Ali, et al. "Nitrous oxide fluxes and nitrifier and denitrifier communites as affected by dry-wet cycles in long term fertilized paddy soils." *Applied Soil Ecology* 125 (2018): 81-87.

[2] Nie, Hongmei, et al. "Spatial Prediction of Soil Moisture Content in Winter Wheat Based on Machine Learning Model." *2018 26th International Conference on Geoinformatics*. IEEE, 2018.

[3] Kaur, Vimaldeep, Gitanjali Sharma, and Chandni Kirpalani. "Agro-potentiality of dairy industry effluent on the characteristics of Oryza sativa L.(Paddy)." *Environmental Technology & Innovation* 12 (2018): 132-147.

[4] Nie, San'an, et al. "Dissolved organic nitrogen distribution in differently fertilized paddy soil profiles: implications for its potential loss." *Agriculture, Ecosystems & Environment* 262 (2018): 58-64.

[5] Razavipour, Teimour, et al. "Azolla (Azolla filiculoides) compost improves grain yield of rice (Oryza sativa L.) under different irrigation regimes." *Agricultural Water Management* 209 (2018): 1-10.

[6] Islam, ARM Towfiqul, Shuang-He Shen, and Shen-Bin Yang. "Predicting design water requirement of winter paddy under climate change condition using frequency analysis in Bangladesh." *Agricultural Water Management* 195 (2018): 58-70.

[7] Tsujimoto, Kumiko, et al. "Estimation of planting date in paddy fields by time-series MODIS data for basin-scale rice production modeling." *Paddy and Water Environment* 17.2 (2019): 83-90.

[8] Cao, Jingjing, et al. "Irrigation scheduling of paddy rice using short-term weather forecast data." *Agricultural water management* 213 (2019): 714-723.

[9] Kumar, P., et al. "Comprehensive evaluation of soil moisture retrieval models under different crop cover types using C-band synthetic aperture radar data." *Geocarto International* 34.9 (2019): 1022-1041.

[10] Salam, Abdus, et al. "Rice straw-and rapeseed residue-derived biochars affect the geochemical fractions and phytoavailability of Cu and Pb to maize in a contaminated soil under different moisture content." *Journal of environmental management* 237 (2019): 5-14.[11] Ding, Changfeng, et al. "Changes in the pH of paddy soils after flooding and drainage: modeling and validation." *Geoderma* 337 (2019): 511-513.

[12] Chen, Haorui, et al. "Development of a waterlogging analysis system for paddy fields in irrigation districts." *Journal of Hydrology* 591 (2020): 125325.

[13] Jiang, Honghua, et al. "CNN feature based graph convolutional network for weed and crop recognition in smart farming." *Computers and Electronics in Agriculture* 174 (2020): 105450.

[14] Baskar, Chanthini, and Manivannan Doraipandian. "Fuzzy Logic-Based Decision Support for Paddy Quality Estimation in Food Godown." *Advances in Electrical and Computer Technologies*. Springer, Singapore, 2020. 279-286.

[15] Chen, Hongping, et al. "The within-field spatial variation in rice grain Cd concentration is determined by soil redox status and pH during grain filling." *Environmental Pollution* 261 (2020): 114151.

[16] Xu, Ying, et al. "Conversion from double-season rice to ratoon rice paddy fields reduces carbon footprint and enhances net ecosystem economic benefit." *Science of The Total Environment* (2021): 152550.

[17] Xu, Peng, et al. "Conversion of winter flooded rice paddy planting to rice-wheat rotation decreased methane emissions during the rice-growing seasons." *Soil and Tillage Research* 198 (2020): 104490.

[18] Maneesha, Akula, Chalumuru Suresh, and B. V. Kiranmayee. "Prediction of rice plant diseases based on soil and weather conditions." *Proceedings of International Conference on Advances in Computer Engineering and Communication Systems*. Springer, Singapore, 2021.

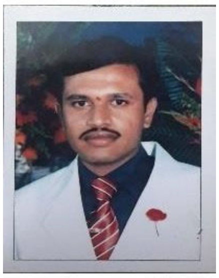
[19] Girinath, N., et al. "Intelligent Irrigation System for Temperature and Moisture Monitoring." *2021 Smart Technologies, Communication and Robotics (STCR)*. IEEE, 2021.

[20] Zhu, A., et al. "Mapping rice paddy distribution using remote sensing by coupling deep learning with phenological characteristics." *Remote Sensing* 13.7 (2021): 1360.

[21]Deb, Mainak, et al. "Paddy Disease Classification Study:A Deep Convolutional Neural Network Approach."*Optical Memory and Neural Networks* 30.4 (2021): 338-357.

[22] Sarkar, Bappa, et al. "Land suitability analysis for paddy crop using GIS-based Fuzzy-AHP (F-AHP) method in Koch Bihar district, West Bengal." *Geocarto International* (2021): 1-27.

BIOGRAPHIES OF AUTHORS



SampathKumar S is Assistant Professor at Kammavari Sangam Institute of Technology, under VTU, Karnataka, India. He is having teaching experience of 17 years, currently perceiving Ph. D degree in the field of Embedded Systems, under VTU, Karnataka, India. His research areas are Agronomy field of Agriculture & Machine Learning.

He has published several papers in refereed presented in National and International Conferences. He has attended several FDPs, conferences and workshops.

Dr.B.N.Manjunatha Reddy

has joined the Global Academy of Technology in the year 2004 and is currently working as a Professor in the Department of Electronics and Communication Engineering. His research and professional career spans about 23 years. He has completed his Ph.D. degree from VTU in the year 2018.His research interests in the area of Low power VLSI and Embedded Systems. He has published several papers in refereed International Journals and presented in National and International Conferences. He has also visited many countries like Dubai, China, Thailand, Abu Dhabi and presented his research work. He is a life member of IETE.

He has organized/attended several FDPs, Conferences and workshops. He also had given several technical talks in many forums.