Received: April 5, 2018

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Deep Learning Based Weighted SOM to Forecast Weather and Crop Prediction for Agriculture Application

Pushpa Mohan[1]* **Kiran Kumari Patil [1]**

[1] *REVA University, Computer Science and Engineering, India*
* Corresponding author's Email: pushpa.mohana@gmail.com

Abstract: One of the most experimentally difficult problem in the world is weather forecasting, which is a basic mechanism in meteorology. Especially in data mining system, there are different information mining strategies are available, for example, K-Means, Artificial Neural Network (ANN) and Support Vector Machine (SVM), etc. These weather predicting strategies are financially high and also very inconsistent for large datasets. To overcome these issues, an effective dimensionality reducing strategy: Self Organizing Map (SOM) is proposed along with Latent Dirichlet Allocation (LDA). The SOM strategy is one of the proper dimensionality reducing strategy to highlight the self-arranging outline. After reducing the measurement, the dimensionality reduced information are used to forecast climate for a reasonable outcome. A reasonable season for an appropriate crop is arranged with the guide of Deep Neural Network (DNN) classification system. This research work depends on finding appropriate information model, which helps in accomplishing high precision and simplification for value forecast. Finally, the experimental outcome shows that the proposed approach improved accuracy in weather and crop prediction up to 7-23% compared to the existing methods.

Keywords: Data mining, Deep neural network, Latent Dirichlet allocation, Self-organizing map, Weather forecasting.

1. Introduction

Data mining is widely used in many fields and it contains various information mining strategies such as ANN, K-Means, SVM, etc. Mainly, the previous researches concentrates on the weather estimation based agriculture application. Agriculture system models are used robust estimation methodologies namely Scaled Conjugate Gradient (SCG) and Broyden Fletcher Goldfarb Shanno (BFGS) Quasi-Newton based on a neural network algorithm. Using this algorithm, predict the soil moisture content to control the farm irrigation [1]. Generally, the water requirement is very essential for agriculture processes, so evapotranspiration estimations are very supportive for farmers. Evapotranspiration estimations give upcoming water specifications clues and hence it is further more helpful for real-time irrigation water management [2].

For developing the smart farming and to reduce the land sizes, labour costs, resources some advanced technologies are applied for getting the best crops. For these improvements, three machine learning algorithms are proposed, namely supervised learning, unsupervised learning and Reinforcement learning algorithms. These algorithms improved the accuracy of artificial intelligence machines with sensor based systems utilized in precision farming [3]. To progress the agriculture business revenues and support the food security and for sustainable agriculture development are depends on the entire result of Climate-Smart Agriculture (CSA) via ensuring the agricultural production system. The CSA sustains severe agriculture that increases indelible feasibility of agriculture and food production [4].

In agricultural fields, developing the crop productivity is the most significant challenges for the farmer [5]. If the person is beginner in the agriculture field. Initially, farmer doesn't know whether which environment is appropriate for which kind of yields, which kind of soil is fit for which crops and which weather is suitable for which crops without knowing following conditions, it is difficult to choose better crops that automatically reduces the income [6, 7]. To overcome these problems, data mining techniques are used. These techniques help in predicting the rainfall, moisture, temperature, wind speed information based on the early years. Based on this prediction, it significantly developed the productivity of the crop [8 - 10].

In this experimental research, for performing crop and weather prediction, an effective technique is introduced, namely SOM-LDA (Weighted-SOM) model. This model is performed through computational modelling, permitting to analyse and decide on various situations related to weather and climate. The main advantage of using SOM is that the data is easily interpreted and understood. The reduction of dimensionality and grid clustering makes it easy to observe similarities in the data. After organizing, the numerical data, which are classified by employing DNN classifier. The major advantages of using DNN classifier includes: it has the ability to implicitly detect complex nonlinear relationships between dependent and independent variables, it can handle large amount of data sets, and it has ability to detect all possible interactions between predictor variables.

This paper is composed as follows. In Section 2, survey several distinctive weather prediction strategies. In section 3, the weighted–SOM is portrayed along with DNN to seek the better weather prediction accuracy. In Section 4, the execution of weight based SOM algorithm for weather prediction is assessed by simulation. Conclusion is made in Section 5.

2. Literature survey

Several techniques were proposed by various authors for crop yield prediction. In this scenario, a brief evaluation of some important contributions to the existing literatures is presented.

J. Mariette, M. Olteanu, and N. Villa-Vialaneix, [11] compared two extensions of the stochastic SOM for dis-similarity data: the first one takes advantage of the online setting in order to maintain a sparse representation of the prototypes at each step of the algorithm, while the second one uses a dimension reduction in a feature space defined by the dis-similarity. In this literature, the dis-similarity data were analysed with topographic maps and also presented a new version of the SOM algorithm, which ensures a sparse representation of the prototypes through online updates. Second, the proposed approach was compared on several benchmarks to a standard dimension reduction technique (K-PCA), which was adapted to large datasets with the Nyström approximation. The developed methodology was not dynamic in forecasting the occasional changes and also it was not eco-friendly.

C. Lennard, and G. Hegerl, [12] developed a supervised scheme: SOM for surface rainfall analysis associated with synoptic circulation. It was investigated for two types of stations in various rainfall areas in South Africa. These synoptic circulations were recognized for winter and summer time as mid-latitude based cyclones, whereas no circulations were related to spring and autumn rainfall. The paper evaluates the capability of SOMs to match the synoptic movers of observed rainfall record, which effectively downscales the large-scale data of synopses to an accurate resolute response of the surface. This method was concluded as relatively inexpensive on the guide of downscaling instrument contrasted with other downscaling methods, particularly local atmosphere models, and it was helpful for breaking down the change and its effects on circulation of atmospheric characteristics. For rainfall prediction, numerous number of attributes were required in SOM to enhance the accuracy of rainfall detection and also several meteorological information were considered to improve the weather observations by specialists.

A.Y. Abdulrahman, T.A. Rahman, I.M. Rafiqul, B.J. Olufeagba, T.A. Abdulrahman, J. Akanni, and S.A.Y. Amuda, [13] updated the necessity for integrating crop-climate structures and also clarified that the integration can assist to overcome current difficulties like mismatch between farmer's requirement and obtainable predictions, risks and time related doubts, task in achieving institutional, financial and political provision, etc. This literature observed the incorporated methods so far established and strongly sustained the integration of crop weather representations for improvement of the predictions. Outlandish coordination of information into the farming decision methodology make a correspondence challenge among scientist and accomplice.

L. Bornn, and J.V. Zidek, [14] described, how spatial dependence was incorporated into statistical models for crop yield prediction. Initially, select biophysically based explanatory variables and partially-determined prior probability distributions. Then, Bayesian model was utilized to increase the modelling flexibility and also for improving the prediction over existing least-squares methods. The model was focused on providing efficient prediction, which stabilize the effects of noisy data. Prior distributions were developed to accommodate the spatial non-stationary arising from distinct between-region differences in agricultural policy and practice. As a result, the model improved the prediction performance relative to existing models and allowed interpretation of climatic effects on the model’s output. These simulation models can't predict the modifications in weather estimating accurately and consequently prompt to incorrect data forecast.

A. Shastry, H.A. Sanjay, and E. Bhanusree, [15] proposed Regression Analysis (RA) to determine the environmental factors and their infliction on crop yield. RA was a multi-variate analysis approach, which analyses the factors and groups them into response variables that helps to obtain a decision. A sample of environmental factors like soil type, crop parameters were considered for a period of 10 years from 1990-2000. This research was extended by considering other factors like minimum support price, cost price index, whole-sale price index etc. The amount of data being generated and stored every day was exponential, especially in crop yield prediction.

J. Scheffel, K. Lindvall, and H.F. Yik, [16] proposed a time-spectral methodology based on Generalized Weighted Residual Method (GWRM) for numerical weather prediction. In this study, comparisons of accuracy and efficiency were carried out for both explicit and implicit time-stepping systems. It was found that the efficiency of GWRM shows better result compared to the existing methods in terms of accuracy. The GWRM has the additional advantage to produce analytical solutions in the form of Chebyshev series expansions. The weather prediction systems were largely driven by prevailing winds, whereas small changes in the wind speed and direction can result in significant changes in weather estimation.

To overcome the above mentioned drawbacks, a combination of SOM and LDA is implemented with multi-objective classifier DNN for enhancing the performance of season and crop prediction.

3. Study area and site description

Weather forecasting is essential in forests management for preventing and controlling wildfires. Also, weather forecasting is very helpful in farm operations such as, to irrigate the crop or not, when to apply fertilizer, whether to start complete harvesting or not. Utility companies (electricity, gas) rely on weather forecasts to expect demand, which is strongly affected by the weather. In such vital arranging of crops and trimming designs, brief period climatic information, both standard and handled, have a fundamental part to play.

Events of inconsistent and unfriendly climate agronomic methodologies must be artificial to adapt the impacts of unpredictable and destructive climate on horticultural generation. For example, delay in the beginning of yield season is countered by utilizing a brief length variety of products and thicker sows and the impacts of ices are prevented by depending on water system or lighting garbage fires. Medium range climate assumptions with a validity period allow the agriculturists to compose and complete proper social operations. The accompanying sort of assumption helps more in operational agro meteorology.

- Preliminary exercises, including land readiness and arrangement of plant material.
- Planting or seeding sowing; Management of products, natural product trees and vines.
- Application of compost, water system; diminishing, topping, weeding; nuisance and infection control; Management of touching frameworks.
- Harvesting, on-ranch post-collect handling.
- Transport of create; Livestock generation exercises (for dairy endeavours, meat frameworks, sheep and other animal’s frameworks).

3.1 Crop yield forecasting

Reliable and timely predict of agricultural service gives essential and precious contributions for appropriate arranging. This allows organizers and leaders to predict the amount to import if there should be an occurrence of deficiency or alternatively, to send out in the event of overflow. Similarly, it allows government to set up essential emergency courses of action for redistribution of sustenance during starvation. In India, there is additionally developing requirement for small scale level arranging and especially the interest for product protection, which expands the requirement for field level yield measurements. Determining agricultural output utilizing space, agro meteorology and land based perceptions includes creating econometric, remote detecting and agromet based model to produce various product yield gauges at national, state and area level beginning with harvest sowing to end of season for 11 noteworthy kharif and ragi crops viz., Rice, Jowar, Maize, Bajra, Jute, Ragi, Cotton, Sugarcane, Groundnut, Rapeseed and Mustard and Wheat.

Furthermore, concentrating on confirmation and liability in occasional determines, one can underline a few more aspects, which are detailed below.

- The confirmation of the utilization and the effect of regular figures must be accomplished.
- The client's criticism towards atmosphere forecasters must be moved forward.
- Reliable and applicable information base for confirmation purposes must be readied and accessible.
- The major probabilistic definition of regular conjectures must be considered.
- Promoting the change of the probabilistic estimate to the event of clients in more discernible and far reaching terms must be empowered (changes regarding dangers, or situations or particular client's activities and related.

3.2 Weather prediction problem characteristics

Climate prediction is a mixed-up system that incorporates various particular fields of mastery. "There are just two strategies to anticipate climate: the exact approach and the dynamical approach". Subsequently, the isolated climate approaches have two fundamental branches: numerical demonstrating and logical preparing of meteorological information.

4. Data set description

The crop, weather and soil dataset for entire Karnataka state (India) is collected from Indian meteorological data and Raithamitra-Karnataka State Department of Agriculture (KSDA). This data set contains the attributes such as crop type, soil type, rain fall rate, temperature, humidity, crop cost, etc. The proposed dataset contains more than 1000 records.

5. Proposed methodology

In this section, a significant method: SOM is employed along with LDA scheme for predicting the suitable season and crop for agriculture application. In numerous researches, the SOM approach is employed as an organizer for reducing the computational cost and reproducible outcome. Additionally, the DNN approach is utilized as a classifier in order to enhance the accuracy of prediction rate. The general block diagram of proposed methodology is denoted in the Fig. 1.

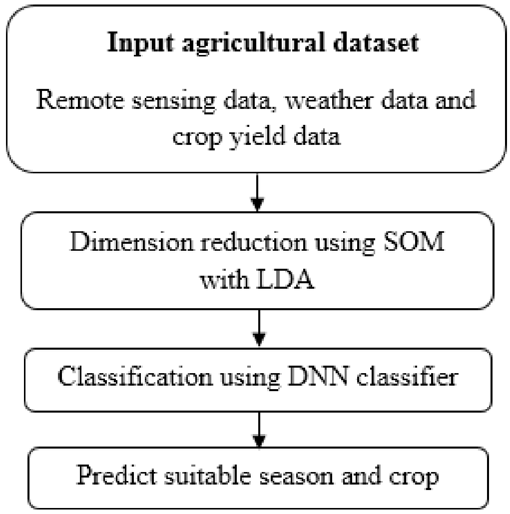

Figure. 1 Block diagram of proposed methodology for weather forecasting

To improve the Euclidean separation work, the element weighting techniques are combined. The proposed WSOM highlighted the weighting strategy, which is the division between class difference and preparation tests inside the class change. As a result, the higher weights are given to the most essential elements, and the lesser weights are offered to the less critical ones. WSOM is an element shrinking approach for DNN classifier that increases the speed of classification time. The classification accuracy of the closest neighbour classifier will improve the quantity of components.

5.1 Dimension reduction using weighted self-organizing map

Self-Organizing Map (SOM) is a type of artificial neural system, which depends on competitive learning procedure. Whereas, SOM contains neuron models that are systematized on a regular two dimensional grid and it has a completely-connected network between the input layer and the neuron layer. The neuron having the smallest distance between input vector and weight vector is named as "winner" neuron. The step by step procedure of weighted-SOM is detailed below,

1) If computational bounds are not exceeding, select an input sample after LDA. The LDA approach is described as a three level Bayesian graphical model, where nodes are presented as random variables and the edges are stands for possible dependencies between the variables.

2) Then, compute the square of the Euclidean distance of weight vectors (w_j) associated with each output node, which is represented in Eq. (1)

$$y_{i,k} = \sum_{k=1}^{n} (x_{i,k} - w_{j,k}(t))^2 \qquad (1)$$

Select output node j^*that has weight vector with minimum value.

3) Whereas, the unit with the minimum Euclidean distance is demonstrated as the winner unit, which it is represented as$j$. Mathematically, the winner unit $j$is stated in Eq. (2).

$$j = arg \min_j\{\|x_n - w_j\|\} \quad (2)$$

4) Since, the self-organizing feature map uses a completive learning rule, the lateral feedback between neurons is often referred as the Gaussian function model. The neighbourhood function around the weighted winner neuron w_j at time t is given in Eq. (3).

$$\Delta(w_j, t) = \exp(-\frac{dj}{\sigma(t)^2}), \text{j= } 1, 2 \dots, n. \quad (3)$$

Where, $dj = \|w_{j(x)} - w_j\|$ is represented as the Euclidean distance between the weighted winning node $w_{j(x)}$ and the corresponding neuron w_j in the lattice, and the parameter$\sigma(t)$ defines an effective width on the feature maps around the winning node. Both $\sigma(t)$ and $\Delta(w_j, t)$ are a monotonically decreasing time varying function. Then the synaptic vector weights are adjusted by all neurons according to the formulation.

5) Finally, update the weights to all nodes within a topological distance by using the weight update rule, which is described in Eq. (4).

$$x = w_j(t+1) = w_j(t) + \eta(t)\Delta(w_j, t)\,[x(t) - w_j(t)], j = 1,2,..n \quad (4)$$

This measurement information set is given as contribution to the following stage for prediction process with the aid of DNN.

5.2 Classification using deep neural network

The DNN typically work as feed forward networks and it is an unsupervised pre-training technique with greedy layer wise training. Here, the data flows from input layer to the output layer without looping function. The major advantage of DNN classifier is, during classification the possibilities of missing value is very low. The DNN technique executes only one layer in unsupervised pre-training stage. The DNN allocates a classification score $f(x)$ during prediction time. To every input data sample $x = [x_1, \dots\dots x_N]$ via a forward pass. Characteristically, f is the function, which involves a sequence of layers for computation, which is represented in Eq. (5).

$$Z_{ij=}x_i w_{ij};\ Z_j = \sum_i Z_{ij} + b_j;\ X_j = g(Z_j) \quad (5)$$

Where, input of the layer is represented as x_i and the output layer is denoted asx_j, w_{ij} are the model parameters and $g(Z_j)$ realizes the mapping or pooling function.

Layer-wise relevance propagation decomposes the classifier output $f(x)$in terms of relevance's r_i attributing to each input component x_i , which contributes to the classification decision described in Eq. (6),

$$f(x) = \sum_i r_i \quad (6)$$

Where, $r_i > 0$ indicates the positive evidence supporting the classification decision and $r_i < 0$ negative evidence of the classification, otherwise neutral evidence. Though, the relevance attribute r_i is calculated using Eq. (7).

$$r_i = \sum_j \frac{zij}{\sum_i zij} \quad (7)$$

The DNN is able to investigate the unknown feature coherences of input. The DNN provides a hierarchical feature learning approach. So, the high level features are derived from the low level features with a greedy layer wise unsupervised pre-training data. Thus, the key objective of DNN is to handle the complicated functions that can represent high level abstraction.

5.2.1. Stacked auto encoder

The deep learning neural network includes multiple layers of sparse auto encoders. Each layer output support to the input of the successive layers. An auto encoder attempts to learn an approximation to the identity function, shown in Eq. (8).

$$\hat{x} = h_{w,b}(x) \approx x \quad (8)$$

The DNN exploits the unsupervised pre-training technique with greedy layer wise training. This technique executes at time one layer in unsupervised pre training, beginning from input to output layer. The first sparse auto encoder (1st hidden layer) is trained on the raw inputs (x) to learn primary features $h^{(1)}$ on the inputs.

Received: April 5, 2018

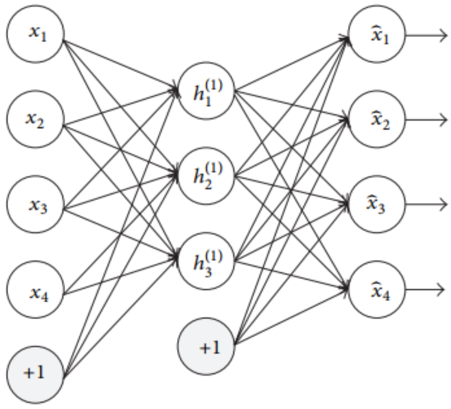

Figure. 2 Structure of an auto encoder

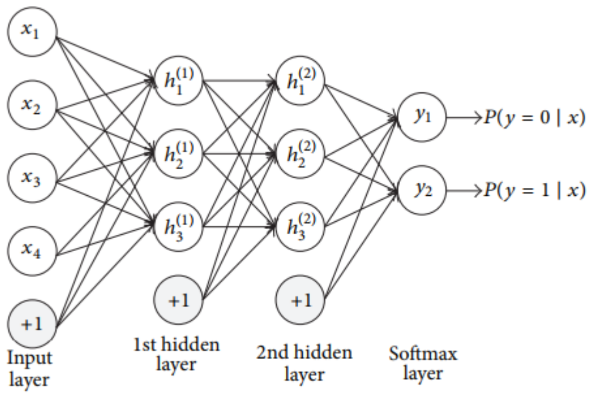

Figure. 3 Stacked auto encoder with softmax classifier

The structure of an auto encoder is depicted in Figure 2. During the pre-training process, all the weight and bias parameters has learned to lessen the cost function. The Figure.3 shows the softmax classifier using auto encoder. The input data use the forward propagation to train sparse auto encoder to attain the basic features.

In the next hidden layer of pre training data, the auto encoder technique calculates its features using the same method from the preceding hidden layers. Equation (9) describes auto encoder,

$$cost = \frac{1}{2n}\sum_{i=1}^{n}(\hat{x}_i - x_i)2 + \beta\sum_{j=1}^{m} KL\ \left(p\middle|\widehat{p}_j\right) + \frac{\lambda}{2}\sum_{i=1}^{n}\sum_{j=1}^{m}\theta_{ij}^2 \quad (9)$$

Where, hidden nodes are represented as m, inputs are n, weight of sparsity penalty as β, the probability of firing activity indicated as $\widehat{p}j$, the sparsity parameter is denoted as ρ, weight delay is represented as λ, KL is Kullback-Leibler divergence function, $\hat{x}_i$ is denoted as the auto encoder hidden unit when the network is given a specific input x and θ is weight of hidden nodes. Using DNN classifier, the weather data is classified successfully. The experimental analysis of weather forecasting and performance calculations of existing and proposed techniques are described in the below sections.

6. Result and discussion

In this scenario, the proposed method was implemented on the Java platform with 3.2 GHz and i5 processor. To predict the weather and crop rate, SOM-LDA in combination with DNN methodology was employed. The performance of the proposed methodology was compared in terms of accuracy, precision, recall, sensitivity and specificity.

6.1 Performance measure

In the proposed method, weather forecasting for agriculture application is evaluated with some of the evaluation metrics. An evaluation metrics is used to evaluate the effectiveness of the agricultural application system to justify theoretical and practical developments of these systems. It consists of a set of measures that follow a common underlying evaluation methodology. The relationship between the input and output variables of a system understand by employing the suitable performance metrics like sensitivity and specificity. The general formula for calculating the sensitivity and specificity of weather forecasting is given in Eqs. (10) and (11).

$$Sensitivity \frac{Number\ of\ TP}{Number\ of\ TP + Number\ of\ FN} \times 100 \quad (10)$$

$$Specificity = \frac{Number\ of\ TN}{Number\ of\ TN + Number\ of\ FP} \times 100 \quad (11)$$

Where, TP is denoted as true positive, FP is denoted as false negative, TN is stated as true negative and FN is stated as false negative.

In addition, accuracy, precision, recall are the suitable evaluation metrics for finding the effectiveness of weather forecasting and crop detection rate. Also, it is the measure of statistical variability and a description of random errors. The general formula of accuracy, precision, and recall for determining weather forecasting and crop detection rate is given in Eqs. (12), (13), and (14).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (12)$$

$$Precision = \frac{TP}{TP + FP} \quad (13)$$

$$Recall = \frac{TP}{TP + FN} \quad (14)$$

Table 1. Comparison of sensitivity and specificity measures of proposed study

Method	Training (%)	Testing (%)	Specificity (%)	Sensitivity (%)
Weighted SOM-DNN	90%	10%	**80.158%**	**82.758%**
	80%	20%	**81.45%**	**83.05%**
	70%	30%	**82.113%**	**83.19%**

6.2 Result of classification evaluation

The performance of the proposed method was estimated based on precision, recall, sensitivity specificity and accuracy value. It was explained in the below tables and also shows the evaluation results for proposed work with validation results. The advanced scheme Weighted-SOM (SOM-LDA) was evaluated in three different combinations of testing and training percentage like 90% training and 10% testing, 80% training and 20% testing, 70% training and 30% testing of collected data. Table 1 shows the sensitivity and specificity measures of proposed study.

The specificity and sensitivity of proposed methodology for 90% training and 10% testing is 80.158% and 82.758%. The 80% training and 20% testing delivers 81.45% of sensitivity and 83.05% of sensitivity. Similarly, the 70% training and 30% testing of collected data delivers, 82.11% of sensitivity and 83.19% of sensitivity. The graphical representation of sensitivity and specificity comparison is specified in Fig. 4.

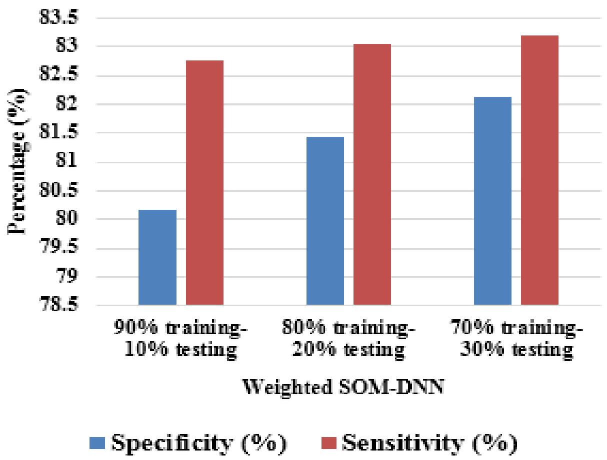

Figure 4. Graphical comparison of specificity and sensitivity measure of proposed method

Table 2. Comparison of precision and recall measures of proposed study

Method	Training (%)	Testing (%)	Precision
Weighted SOM-DNN	90%	10%	**0.827586**
	80%	20%	**0.830508**
	70%	30%	**0.8319**

The Table 2 represents the precision and recall measures for proposed study. The graphical representation of precision and recall value of proposed methodology is specified in the Figure 5. This evaluation metrics confirms that the proposed scheme performs significantly in weather and crop prediction. Table 2 clearly indicates that the proposed method performs effectively in huge weather dataset.

The precision and recall of proposed methodology for 90% training and 10% testing is 0.827586 and 0.88479. The 80% training and 20% testing delivers 0.830508 of precision and 0.8949 of recall. Similarly, the 70% training and 30% testing of collected data delivers, 0.8319 of precision and 0.9 of recall. The graphical representation of precision and recall comparison is stated in the Figure 5.

6.3 Comparative analysis of accuracy

In this section, the proposed method is compared with existing methods: SOM-DNN, SOM-KNN, weighted-SOM-KNN, Random Forests (RF)-Multiple Linear Regressions (MLR) [17] and SOM-Learning Vector Quantization (LVQ) [18]. J.H. Jeong, J.P. Resop, N.D. Mueller, D.H. Fleisher, K. Yun, E.E. Butler, D.J. Timlin, K.M. Shim, J.S. Gerber, V.R. Reddy and S.H. Kim, [17] evaluated a machine-learning method: RF for predicting crop yield responses to climate and biophysical variables at global and regional scales in wheat, maize, and potato in comparison with MLR serving as a benchmark. RF was found highly capable of predicting crop yields and outperformed MLR benchmarks in all performance statistics that were compared. Additionally, P. Mohan, and K.K. Patil [18] proposed an advanced technology, which was the combination of SOM and LVQ methodology. In this study, the prediction accuracy was enhanced by minimizing the within class error among the clusters. The developed approach outcome shows a clear idea about suitable crop cultivation in Mysore region.

In proposed method, the dimension reduction is performed to minimize the data sets, later classification process is done to predict best results, where in the existing method result is directly predicted without use of the dimension reduction. In

Received: April 5, 2018

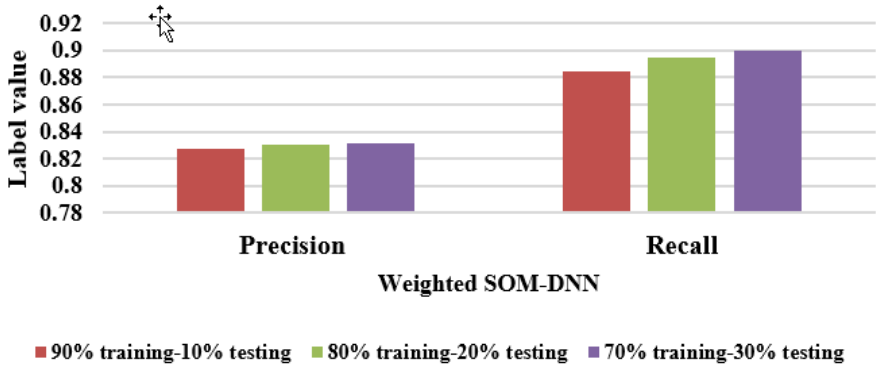

Figure 5. Graphical comparison of precision and recall measure of proposed method

Table 3. Accuracy comparison of existing and proposed work

Method	Training (%)	Testing (%)	Specificity (%)	Sensitivity (%)	Accuracy (%)
SOM-KNN	90%	10%	55.03%	58.80%	57.98%
	80%	20%	52.98%	57.24%	58.99%
	70%	30%	55.07%	56.09%	56.92%
Weighted-SOM-KNN	90%	10%	58.34%	59.98%	58.45%
	80%	20%	57.07%	60.78%	63.32%
	70%	30%	58.35%	60.44%	61.03%
SOM-DNN	90%	10%	61.89%	67.90%	66.90%
	80%	20%	63.59%	65.57%	66.09%
	70%	30%	60.02%	66.66%	68.03%
RF-MLR[17]	90%	10%	71.60%	78.09%	70.12%
	80%	20%	74.42%	72.97%	72.34%
	70%	30%	75.53%	77.74%	76.89%
SOM-LVQ [18]	90%	10%	90.8%	92.3%	67%
	80%	20%	91.2%	92%	65.69%
	70%	30%	89.7%	91%	64.58%
Weighted SOM-DNN	**90%**	**10%**	**80.158%**	**82.758%**	**81.23%**
	80%	**20%**	**81.45%**	**83.05%**	**78.98%**
	70%	**30%**	**82.113%**	**83.19%**	**80.06%**

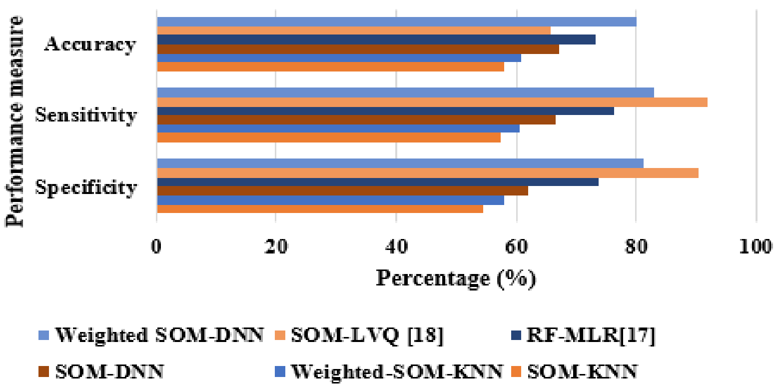

Figure.6 Graphical comparison of accuracy for existing and proposed method

this paper, the proposed weighted-SOM algorithm of DNN classifier is a logical way to reduce the number of features. The weighted-SOM feature map is one of the appropriate dimensionality reduction techniques for DNN classifier, so it can highly maintain topological relationships among samples in a lower dimensional space. Table 3 showed the accuracy comparison of existing and proposed work.

Table 3 shows the performance evaluation of existing methodologies and the proposed method. Figure 6 illustrates the average accuracy, sensitivity and specificity value of proposed method, which is superior compared to existing approaches.

Averagely, the accuracy gap between the proposed and existing methods are 23.04%, 20.093%, 13.09%, 7.02% and 14.34% reduction in accuracy value associated to SOM-DNN, SOM-KNN, weighted-SOM-KNN, RF-MLR [17], and SOM-LVQ [18]. Similarly, the proposed method delivers better performance compared to the existing approaches in terms of sensitivity and specificity.

7. Conclusion

In this experimental research, weather forecasting is processed with the phases like dimension reduction and classification. In the dimensionality reduction process, SOM is utilized along with LDA to perform dimension reduction on acquired dataset. The respective dimensionality reduced data is used for classification by employing DNN. Based on the acquired result, we predict a suitable season with the aid of DNN algorithm. The evaluation result of the proposed method shown better performance than existing methods. On average, the proposed method achieved 80.09% accuracy by utilizing a DNN algorithm. Related to the other approaches for weather and crop prediction, the advanced scheme delivered an effective performance by means of accuracy, sensitivity, specificity, precision and recall, around 7-23% enhancement than the previous methods. In future, for enhancing the agriculture productivity, the weather prediction is further improved by employing a new approach along with suitable crop prediction progression, which is performed on large collection of dataset in a day wise prediction.

References

[1] A.G. Mohapatra and S.K. Lenka, "Neural Network Pattern Classification and Weather Dependent Fuzzy Logic Model for Irrigation Control in WSN Based Precision Agriculture", *Procedia Computer Science*, Vol.78, pp.499-506, 2016.

[2] Y. Luo, S. Traore, X. Lyu, W. Wang, Y. Wang, Y. Xie, X. Jiao, and G. Fipps, "Medium range daily reference evapotranspiration forecasting by using ANN and public weather forecasts", *Water Resources Management*, Vol.29, No.10, pp.3863-3876, 2015.

[3] A. Anitha and D.P. Acharjya, "Crop suitability prediction in Vellore District using rough set on fuzzy approximation space and neural network", *Neural Computing and Applications*, pp.1-18, 2017.

[4] K. Kaur, "Machine Learning: Applications in Indian Agriculture", *International Journal of Advanced Research in Computer and Communication Engineering*, Vol.5, No.4, pp.342-344, 2016.

[5] O.E. Olayide, I.K. Tetteh, and L. Popoola, "Differential impacts of rainfall and irrigation on agricultural production in Nigeria: Any lessons for climate-smart agriculture, *Agricultural Water Management*, Vol.178, pp.30-36, 2016.

[6] C.A. Hölbig, V.L. Machado, and W. Pavan, "Statistical Correction of the Result of Weather Forecast by Applying the Model Output Calibration", *Electronic Notes in Theoretical Computer Science*, Vol.324, pp. 79-90, 2016.

[7] J. Wang, S. Kang, J. Sun, Z. Chen, and N. Song, "Spatial prediction of crop water requirement based on Bayesian maximum entropy and multi-source data", *Transactions of the Chinese Society of Agricultural Engineering*, Vol.33, No.9, pp.99-106, 2017.

[8] H.T. Wong and P.C. Lai, "Weather factors in the short-term forecasting of daily ambulance calls", *International Journal of Biometeorology*, Vol.58, No.5, pp.669-678, 2014.

[9] A. Maleki, M.G. Khajeh, and M.A. Rosen, "Weather forecasting for optimization of a hybrid solar-windepowered reverse osmosis water desalination system using a novel optimizer approach", *Energy*, Vol.114, pp.1120-1134, 2016.

[10] X. Huang, G. Huang, C. Yu, S. Ni, and L. Yu, "A multiple crop model ensemble for improving broad-scale yield prediction using Bayesian model averaging", *Field Crops Research*, Vol.211, pp.114-124, 2017.

[11] J. Mariette, M. Olteanu, and N. Villa-Vialaneix, "Efficient interpretable variants of online SOM for large dissimilarity data", *Neurocomputing*, Vol.225, pp.31-48, 2016.

[12] C. Lennard and G. Hegerl, "Relating changes in synoptic circulation to the surface rainfall response using self-organising maps", *Climate Dynamics*, Vol.44, No.3-4, pp.861-879, 2015.

[13] A.Y. Abdulrahman, T.A. Rahman, I.M. Rafiqul, B.J. Olufeagba, T.A. Abdulrahman, J. Akanni, and S.A.Y. Amuda, "Investigation of the unified rain attenuation prediction method with data from tropical climates", *IEEE Antennas and Wireless Propagation Letters*, Vol.13, pp.1108-1111, 2014.

[14] L. Bornn and J.V. Zidek, "Efficient stabilization of crop yield prediction in the Canadian Prairies", *International Journal of*

Agricultural and Forest Meteorology, Vol.152, pp. 223-232, 2012.

[15] A. Shastry, H.A. Sanjay, and E. Bhanusree, "Prediction of Crop Yield Using Regression Techniques", *International Journal of Soft Computing*, Vol.12. No.2, pp.96-102, 2017.

[16] J. Scheffel, K. Lindvall, and H.F. Yik, "A Time-spectral Approach to Numerical Weather Prediction", *arXiv preprint arXiv*, pp.1704.00561, 2017.

[17] J.H. Jeong, J.P. Resop, N.D. Mueller, D.H. Fleisher, K. Yun, E.E. Butler, D.J. Timlin, K.M. Shim, J.S. Gerber, V.R. Reddy, and S.H. Kim, "Random forests for global and regional crop yield predictions", *PloS One*, Vol.11, No.6, pp. e0156571, 2016.

[18] P. Mohan and K.K. Patil, "Weather and Crop Prediction Using Modified Self Organizing Map for Mysore Region", *International Journal of Intelligent Engineering and Systems*, Vol.11, No.2, pp.192-199, 2018.