



# Wheat crop yield prediction using new activation functions in neural network

Shital H. Bhojani<sup>1,2</sup> · Nirav Bhatt<sup>3</sup>

Received: 12 July 2019 / Accepted: 17 February 2020 / Published online: 16 March 2020  
© Springer-Verlag London Ltd., part of Springer Nature 2020

## Abstract

This research mainly based on multilayer perceptron (MLP) neural networks technique of data mining to forecast the wheat crop yield at the district level. There are many statistical and simulation models available, but the proposed algorithm with new activation function provides promising results in a shorter time with more accuracy. Sigmoid and hyperbolic tangent activation functions are widely used in the neural network. The activation functions play an important role in the neural network learning algorithm. The main objective of the proposed work is to develop an amended MLP neural network with new activation function and revised random weights and bias values for crop yield estimation by using the different weather parameter datasets. MLP model has been tested by existing activation functions and newly created activation functions with different cases including weights and bias values. In this research study, we evaluate the result of different activation functions and recommend some new simple activation functions, named DharaSig, DharaSigm and SHBSig, to improve the performance of neural networks and accurate results. Also, three new activation functions created with little variations in the DharaSig function named DharaSig1, DharaSig2 and DharaSig3. In this research study, variable numbers of hidden layers are tested with the variable number of neurons per hidden layer for the agriculture dataset. Variable values of momentum, seed and learning rate are also used in this study. Experiments show that newly created activation functions provide better results compared to ‘sigmoid’ default neural network activation function for agriculture datasets.

**Keywords** Multilayer perceptron · Forecast · Activation function · Classification · Yield · Neural network · Agriculture

## Abbreviations

DM Data mining  
SMW Standard Meteorological Week

WEKA Waikato Environment for Knowledge Analysis  
MLP Multilayer perceptron  
MAE Mean absolute error  
MSE Mean squared error  
RMSE Root mean squared error  
MAPE Mean absolute percentage error  
RAE Relative absolute error  
RRSE Root relative squared error  
BSS Basic sun shine hours  
MAXT Maximum temperature  
MINT Minimum temperature  
RH1 Morning relative humidity (%)  
RH2 Afternoon relative humidity (%)  
VP1 Morning vapour pressure  
VP2 Afternoon vapour pressure  
AF Activation function  
DA Direction accuracy

Shital H. Bhojani is a Research Scholar, RK University, Rajkot, Gujarat, India and Assistant Professor, AAU, Anand, Gujarat, India; Nirav Bhatt is a Associate Professor, MCA Department, RK University, Rajkot, Gujarat, India.

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s00521-020-04797-8>) contains supplementary material, which is available to authorized users.

✉ Shital H. Bhojani  
shitalbhojani@aau.in  
Nirav Bhatt  
nirav.bhatt@rku.ac.in

<sup>1</sup> RK University, Rajkot, Gujarat, India

<sup>2</sup> AAU, Anand, Gujarat, India

<sup>3</sup> MCA Department, RK University, Rajkot, Gujarat, India

## 1 Introduction

Today's world contains a large collection of data with millions of terabytes, and the size increases over time. These data are unstructured and unorganized and cannot be useful in decision-making. Therefore, these data need to be structured and organized in a manner that can be used easily with a fast-changing life. The capability to manage these millions of terabytes of data in a proper way is a challenge to both researchers and scientists. Agriculture is the most important sector in the country. The government is focusing on doubling the farmer's income by the year 2020. Wheat is the staple food of the country. To meet the need of today's world crop yield requirement, it is necessary to forecast yield in advance for further decision-making. The crop yield forecast is a very difficult task as it is mainly based on weather variables and climatic conditions. The world is facing the problem of global warming, so this factor also plays a major role in crop yield forecasting.

Data mining plays an important role to deal with the large dataset of agriculture and its problem. Agriculture mainly has time series datasets. Many simulation and statistical models are used nowadays to forecast the crop yield at the national/state/district level. To deal with the time series datasets, data mining classification techniques are very much useful to find out the fruitful results. Neural networks are mainly used for classification and prediction. We can apply neural networks for massive datasets to create statistical models of inputted data. Knowledge acquired from the neural network will help the farmers, research scientist, agricultural organizations and also the government officials for generating mankind adequate decisions and with that achieving competing advantage. The importance of this research will be composed (framed) within the context outlined above and its potential to influence planning, crop planting recommendations and decision-making. Specifically, this significance will be demonstrated in several areas of Gujarat in terms of research objectives, scope, benefits and contributions. The result of the proposed algorithm shows that it is importantly useful in amending result of MLP classification technique. In this study, in total seven weather parameters are considered for the wheat crop yield prediction; these are basic sun shine hours (BSS), maximum temperature (MAXT), minimum temperature (MINT), morning relative humidity (RH1), afternoon relative humidity (RH2), morning vapour pressure (VP1) and afternoon vapour pressure (VP2) and the dataset of the yield used for the years 1990–1991 to 2016–2017. Proposed algorithm results were drawn without the overlay of data. However, how to select the range for the random parameters remains an open question. This issue is considered to be one of the most important research gaps in the field of randomized algorithms for training NNs.

It is noticed that scientists and researchers were contributed in the sector of agriculture since decades and are now focusing more on the customized neural network in the different area of agriculture sector. Different prediction models like artificial neural network, multiple linear regression (MLR) and support vector regression were applied on crop yield prediction. When there is a multiple dependent parameters and nonlinear relationship between the input parameters, the artificial neural network is used for the prediction. Aditya et al. developed customized artificial neural networks (C-ANNs) by a varying number of hidden layers, number of neurons in the hidden layer and the learning rate (LR). This C-ANN is applied to the dataset to predict the yield of wheat [1]. WEKA provides a time series analysis platform using which users can develop, evaluate and visualize the forecasting models [2]. Leyla et al. discuss the steps of ANN model development and how information flows through the network that affects the structure of the ANN based on input and output data [3]. The neural network model is composed of artificial neurons that are interconnected, and depending on the network topology, they exchange the actuation signals—and this in the form of an activation transition function [4]. The advance estimates of crop production are needed much before the actual harvest of crops for making various decisions such as pricing, distribution, export and import [5]. ANNs have been widely used in studies of complex time series forecasting, such as weather, energy consumption and financial series [6]. In feedforward neural networks with random hidden nodes (FNNRHN), the learning process does not require iterative tuning of weights [7]. The weights and biases of hidden neurons need not to be adjusted [8]. Machine learning is a field of study that uses the statistics and computer science principles, to create statistical models, used to perform major tasks like predictions and inference [9].

The rest of the paper is organized as follows: Section 2 illustrates work done using different data mining techniques for forecasting. Section 3 contains the detailed information related to study region, data collection and processing, neural network, activation functions, ANN fuzzy models and development–implementation of the MLP algorithm. In Sect. 4, we present the experimental results, the effect of the parameter on yield using Waikato Environment for Knowledge Analysis (WEKA) and performance evaluation of the proposed algorithm. Finally, conclusions are given in Sect. 5.

## 2 Related work

Nowadays data mining entered into each and every sector for extracting useful patterns and knowledge from the large database. Research in the field of agriculture leads to the

use of the data mining. Scientists and researchers from all over the world had a great deal of research on agriculture and generated fruitful results using data mining methods. The research process never ends, and it is continuing over time to have more impressive results and decisions.

A customized artificial neural network (C-ANN) was proposed in [1] for crop yield prediction with the variable number of hidden layers and hidden nodes, i.e. neurons per layer, and different learning rates. Experiments show that the C-ANN provides better results compared to MLR and default ANN based on the  $R^2$  and percentage prediction error. Different experiments were carried out in [2] using the WEKA multilayer perceptron classification algorithm on agriculture datasets. Different WEKA classification algorithms used by [10] evaluated the MAE and DAC for season-based electricity demand forecasting and concluded that support vector machine algorithm provides better forecasting compared to other algorithms. Steps of artificial neural network, i.e. MLP neural network model development, were discussed by [3] for oil production using the sigmoid activation function, one hidden layer and three nodes in this study, and evaluated the results on the base of RMSE. MLP provides better results compared to the regression analysis and concludes that for oil production forecasting MLP model would be beneficial in SOCAR. Brown onion yield level prediction study was focused [4] using the mean of artificial neural network. MLP model was constructed with (1–2–1) configuration and compared with other nonlinear statistical models.

Hyperbolic tangent, log sigmoid and linear activation function are applied by [5] in MLP neural network algorithm on rice and maize crop datasets to predict crop area and crop yield production using the MATLAB environment. After applying the different activation function, the result shows that the difference between the actual values and predicted values is within  $\pm 20\%$  error. New activation function, namely log–log, probit and loglog, are suggested by [6] in a neural network for time series forecasting. The different experiment was conducted using the MATLAB platform using a multilayer perceptron (MLP) algorithm with one hidden layer, with variation in the number of nodes. The evaluation was carried out by the two learning algorithm, namely CFG and LM, and the best model was chosen using the by evaluating the values of MAPE, i.e. mean absolute percentage error for forecasting. The author recommends using the new activation functions for smaller network structure and financial time series datasets. [7, 8] proposed a method for generating random weights and biases and studied the fitting curve of neurons in activation function and how the randomly generated weights and biases affect the approximation capability of the one-dimensional and multidimensional case networks. Hidden layer neurons he applied the sigmoid, Gaussian, Softplus, Sine and Cosine activation functions for fitting the curve.

For FNNRHN, the proposed a process/algorithm for generating the random weights and biases, to set nonlinear fragments of activation functions in the input space region containing data points.

Existing activation functions and its applications are discussed by [9] in the deep neural network. The author revived the different activation functions like sigmoid, Tanh, ReLU, ELUs and its variation and Softmax, Softsign, Softplus, Maxout, Swish and ELISH. Deep learning architecture mostly uses the ReLU in the hidden node and sigmoid in the output layer for prediction. Three different models based on prediction date are constructed by [11] using the MLP ANN for winter wheat yield prediction. Evaluate the model by measuring the different error rates. By comparing these models, WW30\_06 model with the structure 19:19–15–13–1:1 provides the lowest RMSE. [12] Proposed a hybrid ANN algorithm to optimize the weights of the neural network using the Genetic Algorithm (GA). They concluded that the proposed algorithm has increased convergence speed and overcome from the local minima issue. [13] Applied the stepwise regression method and WEKA classification algorithms to predict the wheat crop yield for the different districts of the Gujarat state and concluded that MLP and AR (Additive Regression) provides better results compared to other algorithms.

### 3 Methodology

The proposed research work is mainly focusing on the MLP classification technique with modification in the original algorithm for forecasting the wheat yield at the district level. In the proposed research work, wheat datasets are used to forecast the yield and developed algorithm result values are compared with the values resulted from the WEKA tool. This study mainly focuses on the neural network classification technique of data mining.

#### 3.1 Study region

There are eight agro-climatic zones in the Gujarat state. A total of seven districts and Agro-meteorology observatories are selected for the research purpose as given in Table 1. Datasets are used for the analysis and research purposes from 1990–1991 to 2016–2017.

#### 3.2 Data collection and processing

The actual yield datasets are gathered for the wheat crop from the Directorate of Agriculture, Gandhinagar, and weather datasets are collected from the Agro-meteorology Department, Anand Agricultural University, Anand, Gujarat. Daily datasets were converted to the weekly datasets

**Table 1** List of districts as per the agro-climatic zones

Zone	Region	Selected district	Observatories	No. of parameters used
Zone II	South Gujarat	Bharuch	Bharuch	7
Zone III	Middle Gujarat	Anand	Anand	7
Zone IV	North Gujarat	Sabarkantha	Khedbrahma	7
Zone V	North-west zone	Banaskantha	Sardarkrushinagar	7
Zone VI	North Saurashtra	Bhavnagar	Mahuva	7
Zone VII	South Saurashtra	Junagadh	Junagadh	5
Zone VIII	Bhal and coastal area	Ahmedabad	Arnej	7

using the Standard Meteorological Week (SMW). 25-year datasets are used for building the model. This model is generated especially for the wheat crop yield prediction at the district level, so weather weekly data were used for the period from 44th to 52nd SMW of selected year to 11th SMW of subsequent year of the wheat-growing season.

### 3.3 Neural network

The neural network is basically a graph consisting of an information processing system and various algorithms. The neural network directed graph has input, hidden and sink (output) nodes, which exist in input, hidden and output layers, respectively. The output node determines the predicted value for the application. It usually works only with the numeric data, where  $F = \{V, A\}$ ; vertices  $V = \{1, 2, 3, \dots, n\}$  and arcs  $A = \{(i, j) | 1 \leq i, j \leq n\}$ , with the following:

1.  $V$  contains the input, hidden and output nodes.
2. The vertices are partitioned into layers  $\{1, 2, \dots, k\}$  having 1 as input layer and  $k$  as output layer. Layers 2 to  $k - 1$  are the hidden layers.
3. The arc is the connection between the two nodes. The arc  $(i, j)$  must have the node  $i$  in layer  $h - 1$  and node  $j$  in layer  $h$ .
4. Each arc  $(i, j)$  has associated weight  $W_{ij}$ .
5. Node  $i$  is associated with the activation function  $f_i$ .

An artificial neural network can be of feedforward or feedback network structure. Figure 1a shows the simple structure of the node  $i$  in NN, and Fig. 1b shows the neural network structure for the crop yield forecasting. There are  $k$  input arcs coming from nodes  $\{1, 2, \dots, k\}$  with associated weights and input values  $w_{1i}, \dots, w_{ki}$  and  $x_{1i}, \dots, x_{ki}$ , respectively. The dashed line shows the values that are propagating through the network. The prediction is shown by the  $y_i$  value. Different activation functions ( $f_i$ ) are applied to the input values to flow in the network.

### 3.4 Activation function

Sometimes it is called processing element function or a squashing function. Different activation functions exists

like sigmoid, threshold, Gaussian, etc. When activation function applies to input values, Eq. 1 is formed, which is a sum of products and if bias exists then Eq. 2 is formed. Bias is an integer value associated with each node in a neural network as an extra input. See Eqs. 1 and 2 where  $W_{ij}$  are weight and  $X_{ij}$  are input values associated with the input layer.

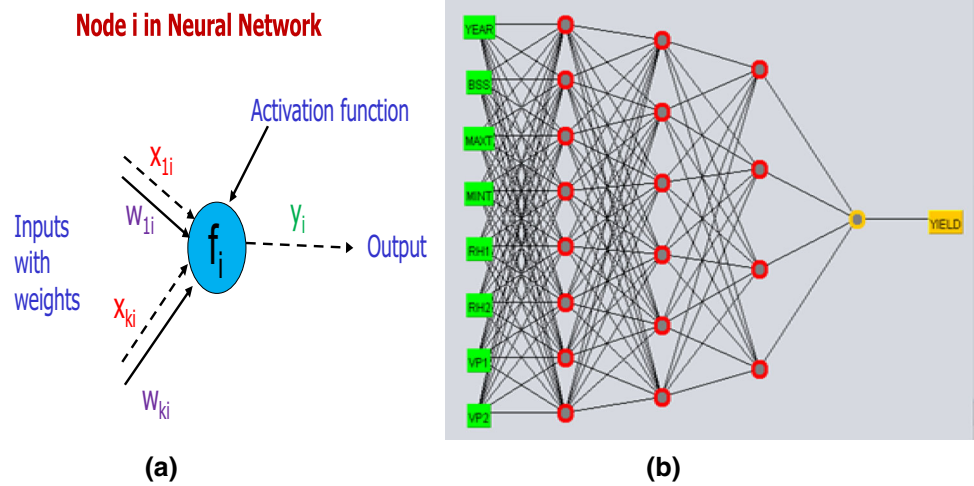
$$S_i = \sum_{j=1}^k (W_{ij}X_{ij}) \quad (1)$$

$$S_i = W_{0i} \sum_{j=1}^k (W_{ij}X_{ij}) \quad (2)$$

### 3.5 Multilayer perceptron (MLP)

MLP is the most common neural network. Neural networks are made up of many artificial neurons, and each neuron has its own weight associated with it. A neuron can have any number of inputs from 1 to  $n$ , where  $n$  (integer value) is the total number of inputs. The inputs are represented as  $x_1, x_2, x_3, \dots, x_n$  with corresponding weights as  $w_1, w_2, w_3, \dots, w_n$  and output as  $a = x_1w_1 + x_2w_2 + x_3w_3 \dots + x_nw_n$ . Neural networks can be used for both classification and numeric prediction. The pseudo-code of algorithm 1 represents the MLP algorithm implemented using WEKA Java files. There are no clear rules as to the 'best' number of hidden layer units. Network design is a trial-and-error process and may affect the accuracy of the resulting trained network. The initial values of the weights may also affect the resulting accuracy. Once a network has been trained and its accuracy is not considered acceptable, it is common to repeat the training process with a different network topology or a different set of initial weights. Cross-validation techniques for accuracy estimation can be used to help decide when an acceptable network has been found. Learning rules of Hebb and delta are used to modify the weights on an input arc. Using the delta rule, we can minimize the error rate  $d_j - Y_i$  at each node.

**Fig. 1** **a** Node in neural network, **b** Neural network with input, output and hidden layers



Algorithm 1:

```

 $w_d$  ;  $d = 1 \dots D$  // Initialize weights for input layer with random values between -1 and 1
 $b$  ; // Initialize bias for hidden layers with random values between -1 and 1
for iteration = 1 . . . Max iteration do
  for all  $(x,y) \in D$  do
     $a = \sum_{i=1}^D w_d x_d + b$  // Apply Activation Function – Default is Sigmoid
    if  $y_a \leq 0$  then
       $w_d = w_d + y x_d$ , for all  $d = 1 \dots D$  // Update weights
       $b = b + y$  // Update bias
    end if
  end for
end for
return  $w_0, w_1, \dots, w_D, b$ 

```

### 3.6 ANN fuzzy models

ANN fuzzy models can also be used for the time series forecasting. Over the past two decades, multitudinous fuzzy time series blueprints have been put forth for agricultural yield production. Bindu [14] focuses on predicting data values on a large spectrum of fuzzy logic computations based on second- and third-degree relationships. Jayaram [15] focused on fuzzy inference systems and fuzzy set theory for crop yield prediction. Hakan [16] discussed some recent applications, developments and improvements in neural network and fuzzy logic to forecast real-life time series problems. Menaka [17] analysed different methods such as artificial neural network, adaptive neuro-fuzzy inference system, fuzzy logic and multilinear regression for crop yield prediction.

### 3.7 Development and implement of MLP algorithm

Currently, many open-source packages are available like ANN packages, R tool, WEKA, Excel Miner, Orange, etc. Java is an object-oriented programming language, easy to learn, platform independent, rich in API, powerful development tools. Therefore, all the experiments were executed using the WEKA open-source Java libraries. By keeping the concept of the WEKA MLP algorithm, a new algorithm is developed specifically for the agriculture crop yield forecasting at a regional level. The command line interface in Java is provided for passing the multiple parameters to perform the multilayer perceptron classification on the selected datasets.



## 4 Experimental results and discussion

A variety of methodologies, algorithms and simulation models have been applied to agriculture datasets for crop yield estimation [10]. The newly generated model is finalized after the different trials on datasets using the MLP algorithm. This model is built with the eight input parameters and their combinations, i.e. input layer, one hidden layer, i.e. the middle layer, variation in hidden nodes on hidden layer and with one output node, i.e. output layer. For forecasting, we have used the ‘three-step ahead’ prediction. This paper mainly focuses on new algorithm development which includes the generation of activation function and random values for input and hidden layer. The experiments were conducted for seven districts of Gujarat state on wheat datasets. All the experiments were executed using the WEKA open-source Java libraries. Multiple sets of different random value initializations were applied for the weights and bias to form a uniform distribution having random values between  $(-1, 1)$ . The neural network was tested by applying the different cases at the input and hidden layers.

Case 1 proposes the neural network having the input weight as zero and the random values on the hidden layer. Case 2 proposes the neural network having the default random values, i.e. weights on the input layer and predefined set of random value array with some arithmetic calculations on the hidden layer. Case 3 proposes the neural network having the default random values, i.e. weights on the input layer and predefined set of random value array on the hidden layer. Case 4 proposes the neural network having an array of predefined values on input and hidden layer according to the default neurons, i.e. 14 neurons on the hidden layer. Case 5 proposes the neural network extending the case 3 with 13 neurons on the hidden layer and case 6 with the 12 neurons on the hidden layer. The maximum neurons for the selected datasets are 14. Experiments show that when we decrease the number of neurons less than 12 the performance of the neural network degrades.

The experimental results achieved by applying new activation functions on each district datasets using the different cases. Some of the existing activation functions are also applied are sigmoid, ReLU, Tanh, Exponential, Reciprocal, Gaussian, Sine, Cosine, Ellicot, Arctan, RadialBasis, Softplus, Leaky ReLU, ELU, Softmax, Cloglog, Cloglogm, Loglog, Sech, Wave, Rootsig and Logsigm. Proposed new activation functions are DharaSig, DharaSigm, SHBSig, DharaSig1, DharaSig2 and DharaSig which improve the performance of the neural network compared to the default sigmoid activation functions. Table 2 shows the proposed activation function formula.

Approximately 500 hundreds of experiments were executed for models and districts, which leads to total thousands of experiments. The result is evaluated on the basis of the values of different error rates like MAE, RMSE, RAE, MAPE, RRSE, RSE and DAC. Apart from this, the observed (actual) and forecasted yield error per cent is also calculated to choose the best-fitted model. The mean absolute percentage error, i.e. MAPE, is a measure of forecast accuracy percentage which plays an important role to find out the best-fitted model in statistics. It is observed that the value of learning rate and momentum is not affecting the result on the output layer for the selected datasets. Research work is carried out by taking the confidence interval at a level of 95%.

### 4.1 Performance evaluations

Different errors rates are calculated in percentage using the formulas given in Table 3, where  $F_i$  = forecasted time series and  $A_i$  = actual time series.

Different errors rate percentages are calculated using the statistical formulas (MAE, MSE, RAE, RMSE, MAPE and RRSE) with forecasted and actual time series data. The MAE is a quantity used to measure predictions of the eventual outcomes. The MAE is an average of the absolute error. RMSE is the square root of the mean of the squares of the values. It squares the errors before they are averaged and RMSE gives a relatively high weight to large errors. The RMSE of a model prediction with respect to the estimated variable is defined as the square root of the mean squared error. The MAPE is a measure of accuracy in a fitted time series value in statistics, which usually expresses accuracy as a percentage. Direction accuracy (DA) is also measured for each and every trial. DA is a measure of the predictive accuracy of a forecasting method in statistics. It compares the upward and downward prediction trends to the actual recognized direction. The default activation function in the neural network is sigmoid. When the sigmoid AF is applied, the actual and predicted yield error rate is between  $-15$  and  $+15$  for the Sabarkantha, Banaskantha, Bhavnagar and Ahmedabad districts. For Bharuch, Anand and Junagadh districts, values are between  $-75$  and  $17$ . So there is a need to amend the algorithm for crop yield forecasting. An experiment shows that the new proposed activation function ‘DharaSig’ and ‘DharaSigm’ producing the best values for case 3, case 4, case 5 and case 6 as described above.

Using the different cases and AF, predicted yields are overestimated and underestimated for selected districts. Case 3 with ‘DharaSig’ AF function (Fig. 2a, b) provides the better results. For Bharuch district, the predicted yield is underestimated by 8.19% for the one-step ahead prediction, i.e. for the year 2015–2016, and overestimated by

**Table 2** Proposed activation functions

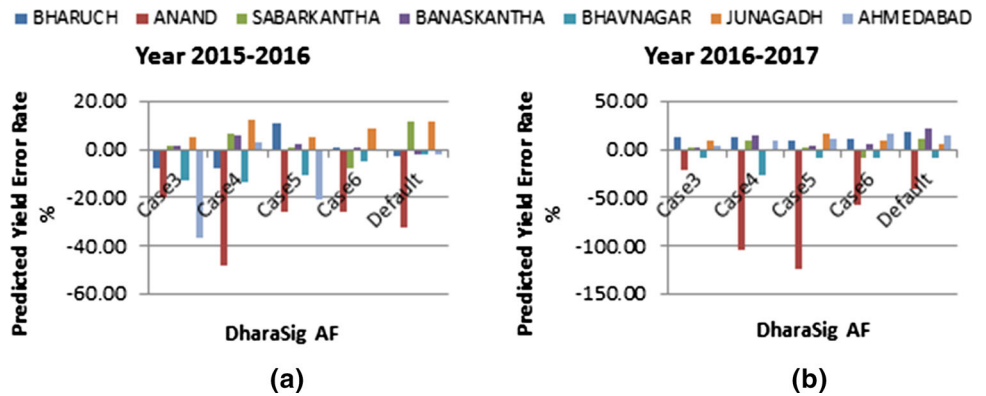
Name	Formula	Name	Formula
DharaSig	$DharaSig = 1 - \exp^{(0.7*f)}$	DharaSig1	$DharaSig1 = 1 - \exp^{(0.1*f)}$
DharaSigm	It is a square of DharaSig, i.e. Let $v = 1 - \exp^{(0.7*f)}$	DharaSig2	$DharaSig2 = 1 - \exp^{(0.2*f)}$
SHBSig	$SHBSig = \frac{1}{\exp^f}$	DharaSig3	Let $v = 1 - \exp^{(0.4*f)}$

Here,  $f = \frac{1}{\exp^{(-2)}}$

**Table 3** Different error rate formulas

Mean Absolute Error (MAE) = $\frac{\sum_{i=1}^N  F_i - A_i }{N}$	Mean Absolute Percentage Error (MAPE) = $\frac{1}{N} \sum_{i=1}^N \frac{ A_i - F_i }{ A_i }$
Mean Squared Error (MSE) = $\frac{1}{N} \sum_{i=1}^N (A_i - F_i)^2$	Relative Absolute Error (RAE) = $\frac{\sum_{i=1}^N  F_i - A_i }{\sum_{i=1}^N  A_i - A_i }$
Root Mean Squared Error (RMSE) = $\sqrt{\frac{\sum_{i=1}^N (F_i - A_i)^2}{N}}$	Root Relative Squared Error (RRSE) = $\sqrt{\frac{\sum_{i=1}^N (F_i - A_i)^2}{\sum_{i=1}^N (A_i - A_i)^2}}$

**Fig. 2** Predicted yield error rate using DharaSig AF **a** for year 2015–2016 and **b** for 2016–2017



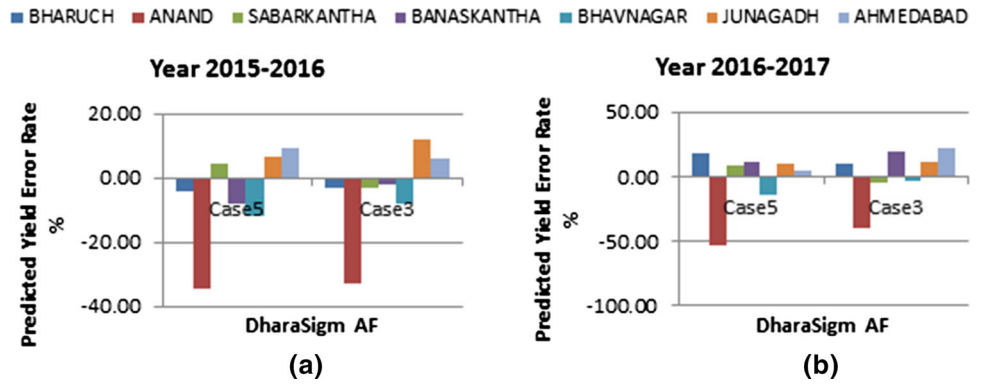
13.46% for the two-step ahead prediction, i.e. for the year 2016–2017. For Anand district, the predicted yield is underestimated by 20.06% and 22.24% for the years 2015–2016 and 2016–2017, respectively. For Sabarkantha district, the predicted yield is overestimated by 1.18% and 1.75% for the year 2015–2016 and 2016–2017, respectively. For Banaskantha district, the predicted yield is overestimated by 1.34% and 1.19% for the year 2015–2016 and 2016–2017, respectively. For Bhavnagar district, the predicted yield is underestimated by 13.14% and 8.90% for the years 2015–2016 and 2016–2017, respectively. For Junagadh district, the predicted yield is overestimated by 5.30% and 8.75% for the year 2015–2016 and 2016–2017, respectively. For Ahmedabad district, the predicted yield is underestimated by 36.49% and overestimated 4.43% for the year 2015–2016 and 2016–2017, respectively. DA is found 100%, i.e. the direction of accuracy is 100% to the upward direction for all the cases and AF's.

Experiments show that the predicted yield error rate is underestimated by 14.51% and overestimated by 18.82% among the selected districts for case 5 with 'DharaSigm' AF function (Fig. 3a, b).

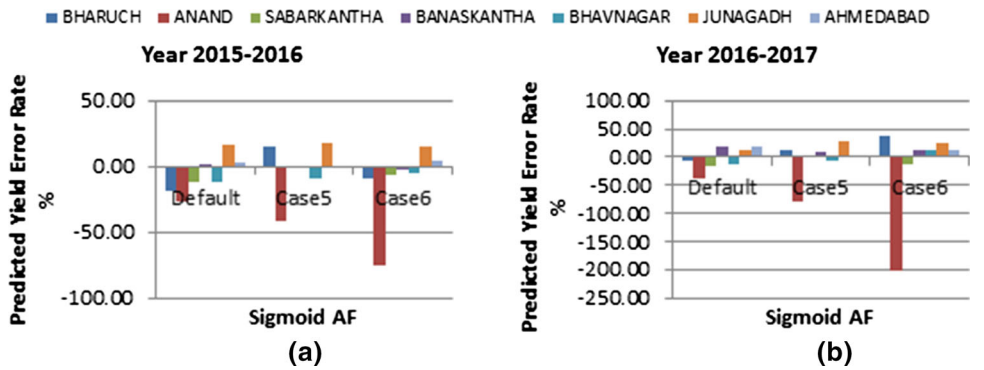
Figure 4a, b shows the result of the 'sigmoid' activation function which is the default in the neural network algorithm.

Experiments show that the predicted yield error rate is underestimated by 8.88% and overestimated by 15.73% among the selected districts for case 6 with 'DharaSig' AF function. Different error values are also calculated to find out the best-fitted model for the agriculture datasets. Activation functions play an important role in the NN hidden layer. Loglog AF also applied to the selected datasets. The predicted yield result of loglog is quite acceptable. Figure 5 shows the comparison of actual yield value with the predicted yield value of different activation functions (Fig. 5a for the year 2015–2016 and Fig. 5b for year 2016–2017 districtwise).

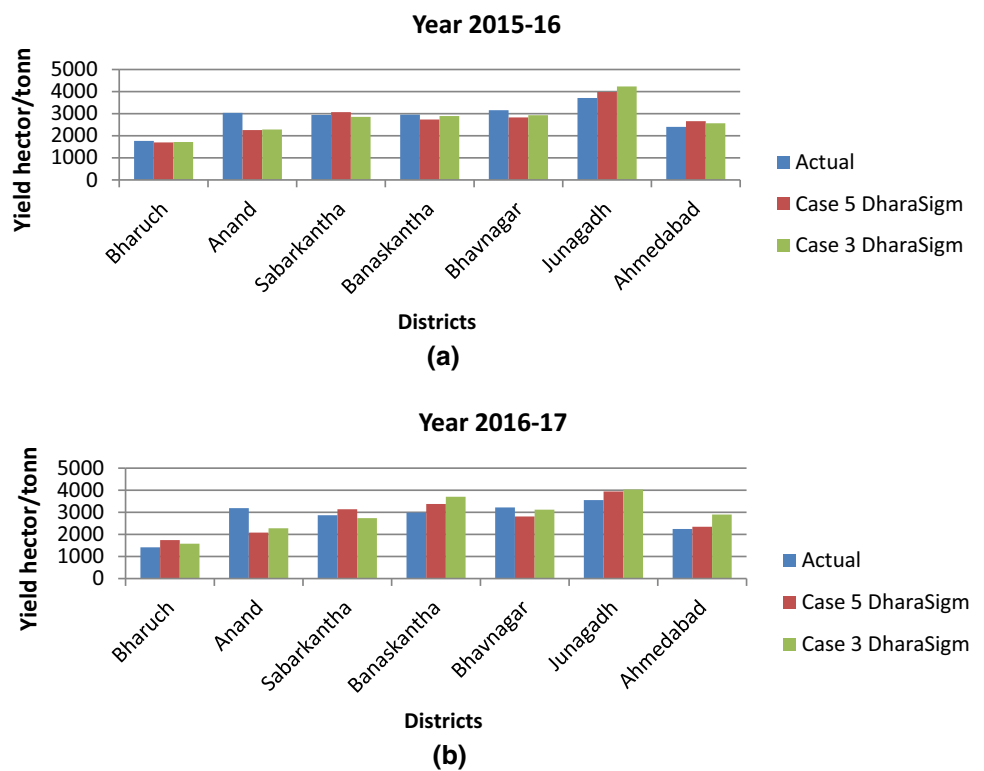
**Fig. 3** Predicted yield error rate using DharaSigM AF **a** for year 2015–2016 and **b** for 2016–2017



**Fig. 4** Predicted yield error rate using sigmoid AF **a** for year 2015–2016 and **b** for 2016–2017



**Fig. 5** Comparison of the actual and predicted yield—**a** for year 2015–2016 and **b** for 2016–2017





**Fig. 6** Comparison of the actual and predicted yield—a for year 2015–2016 and **b** for 2016–2017

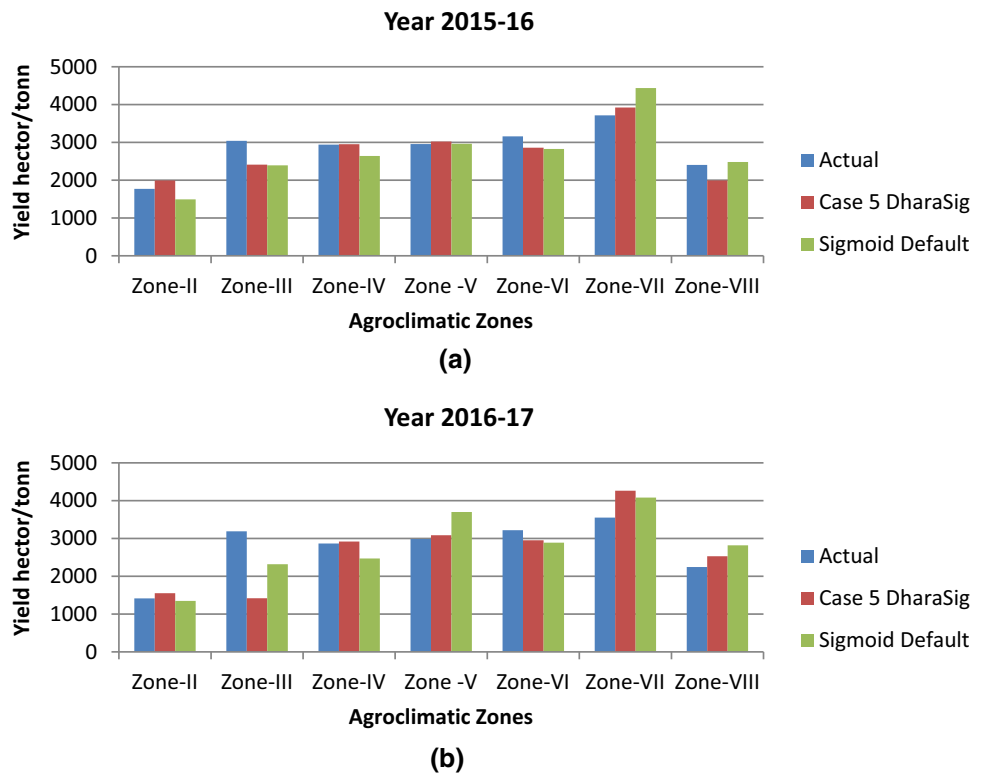


Figure 6 shows the comparison of actual yield value with the predicted yield value of ‘sigmoid’ and new activation function ‘DharaSig’ (Fig. 6a for the year 2015–2016 and Fig. 6b for year 2016–2017 agro-climate zonewise).

Experiments show that for case 3 with ‘DharaSig’ AF, all selected districts except Anand RRSE ranges from 0.0273 to 1.9312, RAE ranges from 0.0257 to 2.1917, MAPE ranges from 0.0061 to 0.2247, RMSE ranges from 0.1119 to 7.1176 and MSE ranges from 0.012 to 50.6608 for the year 2015–2016. RRSE ranges from 0.0268 to 1.9611, RAE ranges from 0.0254 to 2.1131, MAPE ranges from 0.0045 to 0.2253, RMSE ranges from 0.0969 to 7.491 and MSE ranges from 0.0094 to 56.1151 for the year 2016–2017. Case 5 with ‘DharaSig’ AF shows that MAPE ranges from 0 to 0.713 and 0 to 0.759 for the years 2015–2016 and 2016–2017, respectively. Case 6 with ‘DharaSig’ AF shows that MAPE ranges from 0.0062 to 5.7129 and 0.0061 to 4.0576 for the year 2015–2016 and 2016–2017, respectively.

#### 4.2 Effect of the parameter on yield using WEKA

Different classification algorithms applied on time series data of Anand district for the future trend prediction using WEKA ‘Waikato Environment for Knowledge Analysis’ forecast package. For further research multilayer perceptron, a neural network algorithm was selected for yield prediction. Basic and advanced configuration was used for

the crop yield prediction for Anand district. An advanced configuration like lag length, overlay data selection, evaluation metrics and different output options was used. Dataset is preprocessed using the normalization scaling technique and generates new range between 0 and 1 from an existing range of values of different parameters. Anand district dataset was normalized using the min–max normalization technique as shown in Eq. 3.

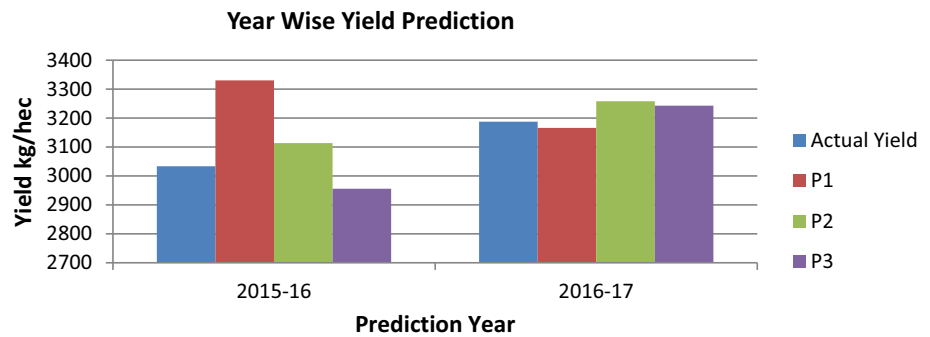
$$v' = \frac{v - \min_A}{\max_A - \min_A} (\text{new\_max}_A - \text{new\_min}_A) + \text{new\_min}_A \tag{3}$$

where  $v$  is the respective original value of the attribute,  $v'$  is min–max normalized value of respective attribute,  $\min_A$  is the respective minimum of value of the attribute and  $\max_A$  is the respective maximum value of the attribute.

Combinations of different variables were applied on the dataset to find the effect of weather parameters on the target field, i.e. crop yield. Overlay of P2 (MAXT, MINT, and BSS) resulting in the best output compared to other parameters for the yield prediction. The combination of P1 (MAXT and MINT) and P3 (MAXT, MINT, VP1 and VP2) provides quite satisfying results too. Figure 7 shows the actual yield value with the predicted yield value for the year 2015–2016 and 2016–2017 for Anand district.

For each case, the amended MLP results are compared with existing functions and this study shows that the

**Fig. 7** Actual and predicted yield value for the year 2015–2016 and 2016–2017



amended MLP provides better results for selected agriculture datasets with a lower prediction error rate.

## 5 Conclusion

Since the last decade, data mining techniques are rapidly growing in the agriculture sector for crop management and the decision-making process. By using the generalized techniques of data mining, researchers are now developing the modified algorithms to focus on some particular aspect or theme. This research paper focuses on the development of the new algorithm for crop yield forecasting. In this research, we proposed three new activation function named DharaSig, DharaSigm and SHBSig for improving the performance of the neural network for agriculture datasets. These newly generated AF results are compared with different cases as described previously in this paper. As compared to the original algorithm of MLP, newly generated algorithm proves the better results with the low RMSE and RAPE and predicted yield forecasting error rate percentage near to 10 for case 3 and DharaSig activation function.  $2n$  ( $n = 7$  weather parameters) combinations were applied on the same datasets, and we concluded that combination of maximum temperature, minimum temperature and basic sunshine parameters results in the best output compared to other parameters combination for Anand district. For further agriculture research, it is recommended to use different crop, weather and soil datasets for the crop yield forecasting at different levels. The same concept can also be applied to the other sector datasets for classification purposes. If large datasets are available then we can remove the noisy data otherwise it needs to be normalized before applying the classification techniques.

**Acknowledgements** I am thankful to the Directorate of Agriculture, Gandhinagar, and Agro-meteorology Department, Anand Agricultural University, Anand, Gujarat, for providing the datasets for my research work.

## Compliance with ethical standards

**Conflict of interest** No conflicts of interest to disclose.

## References

- Aditya Shastry K, Sanjay HA, Deshmukh A (2016) A parameter based customized artificial neural network model for crop yield prediction. *J Artif Intell* 9(1–3):23–32. <https://doi.org/10.3923/jai.2016.23.32>
- Bhojani SH, Bhatt N (2018) Wheat crop yield forecasting using WEKA time series data mining technique. *Int J Sci Res Comput Sci Appl Manag Stud* 7(5):2–3
- Muradkhanli L (2018) Neural networks for prediction of oil production. *ELSEVIER IFAC Papers OnLine* 51(30):415–417
- Stastny J, Konecny V, Trenz O (2011) Agricultural data prediction by means of neural network. *Agric Econ Czech* 57(7):356–361
- Paswan RP, Begum SA, Neog B (2018) Comparison of abilities of different activation functions of artificial neural network to predict crop area and crop production. *Int J Pure Appl Biosci* 6(6):212–220. <https://doi.org/10.18782/2320-7051.7145>
- da Gomes GSS, Ludermir TB, Lima LMMR (2011) Comparison of new activation functions in neural network for forecasting financial time series. *Neural Comput Appl* 20:417–439. <https://doi.org/10.1007/s00521-010-0407-3>
- Dudek G (2019) Generating random weights and biases in feedforward neural networks with random hidden nodes. *Information sciences*. [arXiv:1908.05864](https://arxiv.org/abs/1908.05864) [cs.LG], pp 33–56
- Dudek G (2017) A method of generating random weights and biases in feedforward neural networks with random hidden nodes. [arXiv:1710.04874](https://arxiv.org/abs/1710.04874) [cs.NE]
- Nwankpa C, Ijomah W, Gachagan A, Marshall S (2018) Activation functions: comparison of trends in practice and research for deep learning. [arXiv:1811.03378](https://arxiv.org/abs/1811.03378) [cs.LG]
- Usha TM, Appavu Alias Balamurugan S (2016) Seasonal based electricity demand forecasting using time series analysis. *Circuits Syst* 7:3320–3328
- Niedbała G, Kozłowski R (2019) Application of artificial neural networks for multi-criteria yield prediction of winter wheat. *J Agric Sci Technol* 21:51–61
- Ganatra A, Kosta YP, Panchal G, Gajjar C (2011) Initial classification through back propagation in a neural network following optimization through GA to evaluate the fitness of an algorithm. *Int J Comput Sci Inf Technol* 3(1):98–116

13. Bhojani SH, Bhatt N (2018) Application of data mining technique for wheat crop yield forecasting for districts of Gujarat state. *Int J Sci Res Public* 8(7):302–306
14. Garg B, Aggarwal S, Sokhal J (2018) Crop yield forecasting using fuzzy logic and regression model. *Elsevier Comput Electr Eng* 67:383–403
15. Jayaram MA, Marad N (2012) Fuzzy inference systems for crop yield prediction. *J. Intell. Syst.* 21:363–372. <https://doi.org/10.1515/jisys-2012-0016>
16. Pabuccu H (2017) Time series forecasting with neural network and fuzzy logic. IASSR E-BWN Publication
17. Menaka K, Yuvaraj N (2016) A survey on crop yield prediction models. *Indian J Innov Dev* 5(12)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.