International Journal of Engineering & Technology, 7 (4) (2018) 3204-3208

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET
doi: 10.14419/ijet.v7i4.19767
Research paper

An optimized enhanced Intuitionistic fuzzy cognitive maps for groundnut yield prediction

Malarkodi.K. P [1] *, Dr. Arthi.K [2]

[1] *Part time Research scholar, Research and Development Center, Bharathiyar University, Coimbatore, Tamilnadu*
[2] *Assistant Professor, Department of Computer Science, Government Arts College, Coimbatore, Tamilnadu*
**Corresponding author E-mail: malarkodiphd2017@gmail.com*

Abstract

Over the past decades, several types of crop yield prediction systems using different kinds of data mining algorithms have been developed in agriculture that supports cultivators to analyze the yield productivity. Among those techniques, Fuzzy Cognitive Map (FCM) based crop yield prediction has better efficiency, flexibility and ability to predict yield productivity. However, the performance of FCM was degraded due to some missing input data. Hence in this article, Intuitionistic Fuzzy Cognitive Map (IFCM) is initially used to improve the groundnut yield prediction with the aid of weather and soil parameters. The IFCM is built by considering the expert's hesitancy in the computation of the causal relations between the concepts of a groundnut yield. On the other hand, the learning rate and stability of the IFCM are less due to fixed parameter based weight adaptation. As a result, a supervised multistep learning using the gradient method is proposed for enhancing weight adaptation of IFCM. The enhanced IFCM (EIFCM) estimate the current value of the weight matrix elements from the previous estimation history. Moreover, the learning parameters of the gradient method utilized in EIFCM are optimized by using Self-Organizing Migration Algorithm (SOMA) to reduce the iteration of the weight update. The experimental results prove the efficiency of the proposed OEIFCM in crop yield prediction in terms of accuracy, precision and recall.

Keywords: *Crop Yield Prediction; FCM; IFCM; Gradient Method; SOMA.*

1. Introduction

Generally, agriculture plays an essential role in every country to increase the financial system. As a result, agriculture management has developed by means of different ways like a prediction of crop yield, crop diseases, soil moisture, etc. Such agriculture monitoring or prediction is achieved based on the different characteristics of the atmosphere, crop and soil conditions during the specific time duration [1]. To achieve such predictions, different data mining techniques have been designed [2].

Normally, data mining technique has the aim of extracting the knowledge from an existing dataset and transforming it into a human understandable pattern for advance use [3]. It is the process of analyzing the data from different perceptions and summarizing it into useful information. Most of the research work in agriculture focuses on biological mechanisms for identifying the crop growth and improving its yield. Additionally, it helps decision makers or cultivators to take a decision in respect to the surplus or insufficiency production conditions and permit timely import and export decisions.

Among different crop yield prediction systems, a yield prediction model [4] was proposed based on the dynamic influence graph of the FCM algorithm. In this approach, a data-driven non-linear FCM learning method was proposed for categorizing the yield in apples in which some of decision-making algorithms were described. The proposed FCM has nodes which are connected by directed edges where the nodes were represented as the main soil factors affecting the yield like soil temperature, potassium (K), calcium (Ca), phosphorus (P) and Organic Matter (OM). The cause-effect relationship between the soil properties and yield was defined by the directed edges. However, this approach can be degraded due to some missing input data.

Hence in this paper, the IFCM model is proposed for groundnut yield prediction by considering both weather and soil parameters. IFCM is mostly required to predict the crop yield since it is computationally efficient and it achieves proper decision using more information. In this model, IFCM is generated to learn the input parameters by considering the expert's hesitancy in the computation of the causal relations between the concepts of a domain. In IFCM, fuzzy sets are generalized that their elements are characterized by both membership and non-membership values. However, the decision making of IFCM is still very limited because of its weight adaptation. To improve the weight adaptation, EIFCM is proposed that introduces a supervised multistep learning using the gradient method to learn the IFCM. In this algorithm, the current weight matrix elements are estimated according to the previous estimation. Furthermore, the learning time or computation time complexity is reduced by optimizing the learning parameters which are used in weight adaptation. Thus, an OEIFCM model is proposed based on the optimization algorithm namely SOMA. Thus, the OEIFCM reduces iteration and computational complexity of weight updating while increases prediction accuracy.

The rest of the article is structured as follows: Section 2 presents the previous researches related to crop yield prediction. Section 3 explains the proposed crop yield prediction. Section 4 illustrates the performance effectiveness of the proposed algorithms and Section 5 concludes the research work.

2. Literature survey

A parameter-based model [5] was proposed for crop yield prediction. Here, the crop yield was determined by attributes. The yield of wheat was predicted by using Fuzzy Logic (FL), Adaptive Neuro-Fuzzy Inference System (ANFIS) and Multiple Linear Regression (MLR) techniques. The prediction was achieved by considering biomass, extractable soil water, radiation and rain as different input parameters. The database was pre-processed by means of eliminating outliers, redundant, inconsistent and missing values. The yield of wheat was more accurately predicted by the ANFIS method however the Mean Square Error (MSE) value of the method was slightly high.

A crop yield prediction system [6] was proposed based on the analysis of soil behavior. In this system, the types of the analyzed soil datasets were predicted by using data mining techniques. The predicted type was used for identifying the yielding of crops. Here, the problem of crop yield prediction was formalized as a classification rule in which Naive Bayes (NB) and K-Nearest Neighbor (KNN) techniques were applied to the soil dataset taken from the soil testing laboratory in Jabalpur, Madhya Pradesh. However, only a small dataset was utilized in this analysis due to the occurrence of a few complexities.

Crop yield prediction [7] was proposed by using an Optimal Neural Network (ONN) classifier in spatial data mining. In this system, three major processes were performed such as pre-processing, feature selection and prediction. Pre-processing was used for generating a better model and feature selection process was achieved based on the Multilinear Principal Component Analysis (MPCA). Finally, an ONN classifier was applied to predict the crop yield and realize the precision agriculture. However, the prediction accuracy was not efficiently analyzed.

A sugarcane yield prediction technique [8] was proposed based on the hybrid method. A novel hybrid technique based on fuzzy cognitive map learning algorithm was proposed for sugarcane yield classification. The hybrid algorithm (FCM-DDNHL-GA) was developed by integrating the features of data drivel nonlinear Hebbian learning algorithm and genetic algorithm. This algorithm was improved for different soil and weather features. The accuracy of the classification and inference abilities of this hybrid learning algorithm was evaluated and compared to the machine learning algorithms. However, the performance of evolutionary computation for classification was required further improvement for agricultural monitoring applications.

Ensemble machine learning model [9] was proposed for crop yield prediction. In this model, an AdaBoost algorithm was ensemble with Support Vector Machine (SVM) and Naive Bayes as AdaSVM and AdaNaive. This ensemble model was used for predicting the crop production over a given time period. Initially, historical crop production data and climatic data were collected and combined together. Then, the model was built by classifying the number of input data based on the SVM and naive Bayes algorithms. This process was improved by using this ensemble classification model. However, this model requires an improvement on prediction performance.

A prediction of tea production [10] was proposed by using clustering and association rule mining in Kenya. In this system, the K-means clustering method was used to cluster the dataset based on the specific attributes. Then, the association rule mining was applied to establish an association between the variables based on their occurrence rate. Finally, the prediction of production was achieved by assigning weights to those variables. However, it has a limitation that k-means clustering was depending on the initial cluster centroid estimation.

A robust and novel regression-based fuzzy time series algorithm [11] was proposed for predicting the rice yield. In this algorithm, fuzzy time series approach deals with qualm, obfuscation, veracity, and spuriousness the various facets of the fuzzy contexture. Moreover, frequency-based partitioning was used as the partition of discourse and actual production as the universe of discourse. Then, fuzzy logical relationships of different degrees were executed for effectuating the fuzzification process. Additionally, regression analysis model was applied for achieving the defuzzification process. However, in this algorithm, the MSE value was high.

A crop recommendation system [12] was proposed by using the machine learning algorithm to enhance the crop yield. In this system, an accurate prediction model was built by using a voting model. Initially, different parameters related to crop, soil and environment were collected from soil testing laboratory dataset. Then, the obtained data was given to the recommendation system which has ensemble model with majority voting technique using SVM and ANN as learners for recommending a crop for a specific parameter. However, this system utilizes only a few numbers of attributes.

3. Proposed methodology

In this section, the proposed OEIFCM algorithm for predicting groundnut yield is explained in brief. At first, both weather and soil parameters are gathered and given as input training data to the IFCM system. In this system, the obtained parameters are learned according to the weight values provided by the experts. In IFCM, the fuzzy set elements are characterized not only by a membership value but also by a non-membership value. This IFCM has the ability to consider the degree of hesitancy in the membership value. Thus, a factor of hesitancy is introduced in the cause-effect relations among the concepts of the FCM. Also, the weight adaptation of IFCM is achieved based on the learning method i.e., supervised based gradient method that estimates the current value of weight matrix elements using the previous estimations. Furthermore, the learning parameters used in weight adaptation process are optimized by using SOMA efficiently.

3.1. Enhanced intuitionistic fuzzy cognitive maps (EIFCM)

Based A fuzzy set of IFCM is denoted as,

$$s = \{\langle x, \mu_s(x), \gamma_s(x)\rangle | x \in E\} \quad (1)$$

Here, $\mu_s: E \to [0,1]$ and $\gamma_s: E \to [0,1]$ are the membership and non-membership degree of the element $x \in E$ to the set $s \subset E$. For each element $x \in E$, it holds that $0 \leq \mu_s \leq 1, 0 \leq \gamma_s \leq 1$ and,

$$0 \leq \mu_s(x) + \gamma_s(x) \leq 1 \quad (2)$$

For each $x \in E$, if $\gamma_s(x) = 1 - \mu_s(x)$, s represents a fuzzy set. According to the IFCM model, the cause-effect relations between two concepts C_i and $C_j, i, j = 1, \dots, N$ are defined by both their influence and degree of hesitates. The hesitancy of the element $x \in E$ to the set $s \in E$ is denoted by a fuzzy set $\left(\widetilde{H}_n\right)_{ij}$, on [0,1], from $\widetilde{H} = \{\widetilde{H}_n\}, n = 1, \dots, h$. A subset is selected as,

$$\widetilde{\Omega} \subseteq \widetilde{I} \times \widetilde{H} \quad (3)$$

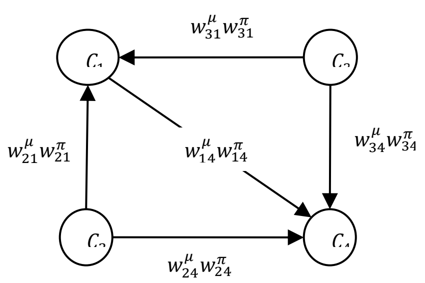
Where

$$\widetilde{I} \times \widetilde{H} = \{(\widetilde{I}_1, \widetilde{H}_1), (\widetilde{I}_1, \widetilde{H}_2), \dots, (\widetilde{I}_2, \widetilde{H}_1), (\widetilde{I}_2, \widetilde{H}_2), \dots, (\widetilde{I}_g, \widetilde{H}_h)\} \quad (4)$$

In equation (4), $\widetilde{I}$ refers the influence of concepts. A set Ω of IFS is built based on the following equation,

$$\widetilde{\Omega} = \{\langle x, \mu_{\Omega_n}(x), \gamma_{\Omega_n}(x)\rangle | x \in [-1,1]\}, n = 1, \dots, w \leq m.p \quad (5)$$

In equation (5), $\mu_{\Omega_n}(x) = \mu_{\widetilde{I}_m}(x), m = 1, \dots, g$ refers to a membership function, $\gamma_{\Omega_n}(x) = 1 - \mu_{\widetilde{I}_m}(x) - \mu_{\widetilde{H}_p}(x), p = 1, \dots, h$ refers to a non-membership function and $\pi_{\Omega_n}(x) = \mu_{\widetilde{H}_p}(x)$ denotes the hesitancy function. The representation of the IFCM model is shown in Figure 1.



$\boldsymbol{C_1 - C_4}$: Concepts $\boldsymbol{w_{ji}^{\pi}}$: Hesitancy weight

$\boldsymbol{w_{ji}^{\mu}}$: Influence weight

Fig. 1: Representation of IFCM Model with Four Nodes.

The considered Concepts of IFCM model and their linguistic values are given in Table 1.

Table 1: Concepts of the IFCM Model and the Values

Concepts	Linguistic Values
C_1: Wind	"Low", "Medium", "High"
C_2: Humidity	"Low", "Medium", "High"
C_3: Air Temperature	"Low", "Medium", "High"
C_4: S.S Temperature	"Low", "Medium", "High"
C_5: Soil Temperature	"Low", "Medium", "High"
C_6: Potassium (K)	"Low", "Medium", "High"
C_7: Calcium (Ca)	"Low", "Medium", "High"
C_8: Zinc (Zn)	"Low", "Medium", "High"
C_9: Organic Matter (OM)	"Low", "Medium", "High"

Consider the hesitancy has a negative impact on the cause-effect relations among the concepts. Hence, the value of each node in each state vector $s_i^t \in [0,1], i = 1, \dots, N$ is represented as follows,

$$s_i^{t+1} = f\left(s_i^t + \sum_{\substack{j=1 \\ j\neq i}}^{N} s_j^t \cdot w_{ji}^{\mu} \cdot \left(1 - w_{ji}^{\pi}\right)\right) \tag{6}$$

Here, w_{ji} is the weight value of the edge directed from node j to node i, $w_{ji}^{\mu} \in [-1,1]$ and $w_{ji}^{\pi} \in [0,1]$ are influence weight and hesitancy weight related to the edge directed from node j to node i. Also, the weight factor $w_{ji}^{\mu} \cdot \left(1 - w_{ji}^{\pi}\right)$ preserves the sign of the influence and considers a zero value when two concepts are not related or the hesitancy weight is equal to unity. If the hesitancy value is zero, then the above equation (6) will depend only on the influence weight. As a result, only weight values of IFCM are chosen by a multistep supervised learning based on gradient method that enhances the learning process of IFCM and also improves the stability with an increased rate of learning. The current value of both influence and hesitancy weight matrix elements is estimated based on the previous estimations. For influence weight matrix $w_{j,i}^{\mu}$ and hesitancy weight matrix $w_{j,i}^{\pi}$, the gradient method is described by the following equations:

$$w_{j,i}^{\mu}(t+1) = P_{[-1,1]}\left(\sum_{k=0}^{m_1} \alpha_k \cdot w_{j,i}^{\mu}(t-k) + \sum_{l=0}^{m_2}\left(\beta_l \cdot \eta_l(t) \cdot \left(R_i(t-l) - s_i(t-1)\right) \cdot y_{j,i}(t-l)\right)\right) \tag{7}$$

$$w_{j,i}^{\pi}(t+1) = P_{[0,1]}\left(\sum_{k=0}^{m_1} \alpha_k \cdot w_{j,i}^{\pi}(t-k) + \sum_{l=0}^{m_2}\left(\beta_l \cdot \eta_l(t) \cdot \left(R_i(t-l) - s_i(t-1)\right) \cdot y_{j,i}(t-l)\right)\right) \tag{8}$$

In equations (8) & (9), $\alpha_k, \beta_l, \eta_l$ are learning parameters which are calculated by using a trial-and-error method, k and l are the number of steps using in the method, t is the learning time ($t = 0,1, \dots, T$) where T is end time of learning, $y_{j,i}(t)$ is a sensitivity function, $P_{[-1,1]}(x)$ and $P_{[0,1]}$ are a design operator for the set $[-1,1]$ and $[0,1]$. Those design operators are defined by an exemplary relation:

$$P_{[-1,1]}(x) = \begin{cases} 1, & x \geq 1 \\ x, & -1 < x < 1 \\ -1, & x \leq -1 \end{cases} \tag{9}$$

$$P_{[0,1]}(x) = \begin{cases} 1, & x \geq 1 \\ x, & 0 < x < 1 \\ 0, & x \leq 0 \end{cases} \tag{10}$$

Sensitivity function $y_{j,i}(t)$ is defined as,

$$y_{j,i}(t) = \left(y_{j,i}(t) + s_j(t)\right) \cdot f'\left(s_i^t + \sum_{\substack{j=1 \\ j\neq i}}^{N} s_j^t \cdot w_{ji}^{\mu} \cdot \left(1 - w_{ji}^{\pi}\right)\right) \tag{11}$$

In equation (12), $f'(x)$ refers to a derivative of the stabilizing function. The termination criterion for the gradient method (8) & (9) is expressed by the following formula:

$$Z(t) = \frac{1}{n}\sum_{i=1}^{n}\left(R_i(t) - s_i(t)\right)^2 < e \tag{12}$$

In equation (13), $s_i(t)$ denotes the value of i^{th} concept $R_i(t)$ denotes the reference value of i^{th} concept and e refers to the level of error tolerance. Also, the learning parameters $\alpha_k, \beta_l, \eta_l$ should satisfy the following conditions to achieve the convergence of the multistep gradient method:

$$\sum_{k=0}^{m_1} \alpha_k = 1 \tag{13}$$

$$0 < \eta_l < 1 \tag{14}$$

$$\eta_l(t) = \frac{1}{\lambda_l + t}, \lambda_l > 0 \tag{15}$$

$$\beta_l \geq 0 \tag{16}$$

According to this method, the current value of influence and hesitancy weight matrix is estimated using previously estimated values that avoid instability issue of weight adaptation process. Moreover, the learning parameters used in this estimation process such as $\alpha_k, \beta_l, \eta_l$ are optimized by SOMA to reduce the number of iterations i.e., reduce the computation time efficiently.

3.2. Optimized enhanced intuitionistic fuzzy cognitive maps (OEIFCM)

In this system, SOMA is applied to optimize the learning parameters $\alpha_k, \beta_l, \eta_l$ which are used to estimate the current value of influence and hesitancy weight matrix elements rapidly. This algorithm has different parameters such as Step, PathLength, PopSize, PRT and the fitness function. Initially, the population of individuals (PopSize) is generated randomly. Each learning parameter for each individual is selected randomly. Each individual from $PopSize$ is evaluated by the fitness function and the individual with the highest fitness is chosen as leader (L) for the current migration loop. After that, all other individuals are migrated to the leader based on $Step$ value.

After each migration, each individual is evaluated using the fitness function. This migration continues until a new location defined by PathLength has been reached. The new location $x_{i,j}$ is computed by using the following equation:

$$x_{i,j(new)} = x_{i,j,start(ML)} + \left(x_{L,j(ML)} - x_{i,j,start(ML)}\right)t \cdot PRTVector_j \tag{17}$$

Here, $t \in < 0, Step, PathLength >$ and ML refers the actual migration loop.

Before an individual starts migration towards the leader, a random number ($rand$) is generated and then compared with PRT. If the generated $rand$ is larger than PRT, then the associated component of the individual is set to 0 by means of PRTVector. This process is continued until the maximum number of migration loops is reached and the best solutions i.e., optimal learning parameters are selected. The following Table 2 gives some examples of the obtained influence and hesitancy weight values.

Table 2: Influence and Hesitancy Weight Values

Cause-Effect Relation	Linguistic Weights [Influence (Hesitancy)]			Real Weights	
	First	Second	Third	w_{ji}^{μ}	w_{ji}^{π}
$C_1 - C_4$	Low (V.Low)	Medium (V.Low)	Low (V.Low)	0.31	0.11
$C_2 - C_5$	Medium (V.Low)	High (Low)	High (V.Low)	0.57	0.16
$C_3 - C_1$	V.Low (V.Low)	V.Medium (Low)	V.Medium (Low)	-0.43	0.18
$C_7 - C_4$	High (V.Low)	High (V.Low)	High (Low)	0.61	0.15
$C_9 - C_6$	V.Low (V.Low)	V.Low (V.Low)	V.Low (V.Low)	-0.29	0.09
$C_5 - C_7$	V.High (V.Low)	V.High (Low)	V.High (V.Low)	0.76	0.15
$C_6 - C_8$	High (V.Low)	High (V.Low)	Medium (V.Low)	0.58	0.11
$C_3 - C_5$	Medium (Low)	Low (V.Low)	Medium (Medium)	0.41	0.25
$C_7 - C_9$	V.Low (V.Low)	V.Low (Low)	V.Medium (V.Low)	-0.39	0.12

Also, Table 3 presents the decision concept's values per iteration obtained from each of the compared models until they reach convergence in seven cases.

Table 3: Representative Decision-Making Examples Using the Compared Models

Case	Given Concepts	IFCM	EIFCM	OEIFCM	Expert's Decision
1	3	High	Low	Low	Low
2	8	High	Medium	High	High
3	5	Medium	Low	Medium	Medium
4	2	Medium	High	Medium	Medium
5	7	Low	V.High	High	High
6	4	Medium	High	V. High	High
7	9	V.Low	Medium	Low	Low

4. Result and discussion

In this section, the performance efficiency of the proposed OEIFCM model is evaluated by using MATLAB 2018a and compared with the other crop yield prediction models such as EIFCM and IFCM in terms of precision, recall, f-measure and accuracy. In this experiment, weather and soil datasets are gathered. The weather dataset includes wind, humidity, air temperature and S.S temperature. The soil dataset consists of soil temperature, potassium, calcium, OM, zinc.

4.1. Precision

It is calculated based on the yield prediction at True Positive (TP) and False Positive (FP) rates.

$$Precision = \frac{TP}{TP+FP} \quad (18)$$

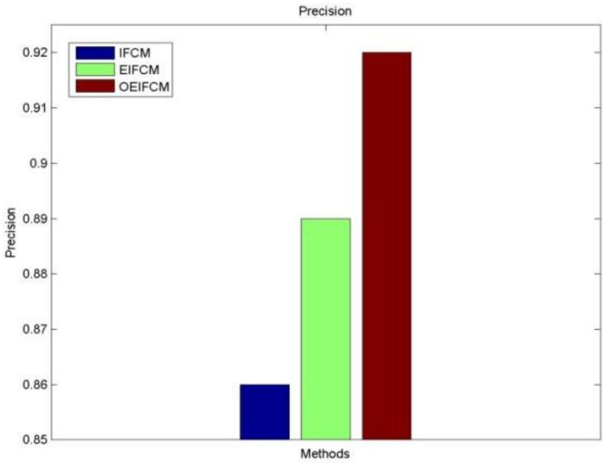


Fig. 2: Comparison of Precision.

Figure 2 shows the comparison of proposed and existing algorithms in terms of precision. The precision of the proposed OEIFCM algorithm is 3.37% higher than EIFCM and 6.98% higher than IFCM. From this analysis, it is observed that the proposed OEIFCM algorithm has better precision than the other algorithms.

4.2. Recall

It is calculated based on the yield prediction at TP and False Negative (FN) rates.

$$Recall = \frac{TP}{TP+FN} \quad (19)$$

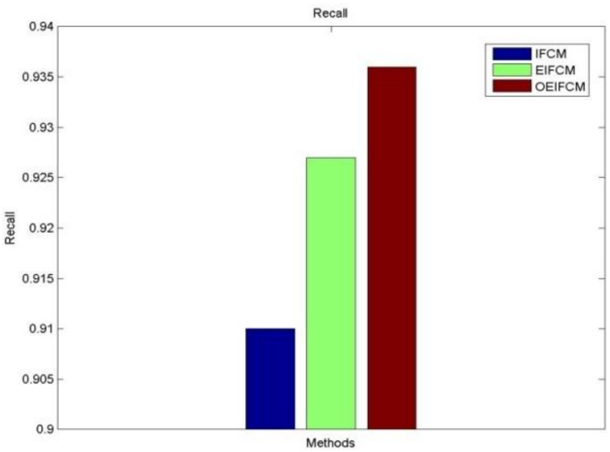


Fig. 3: Comparison of Recall.

Figure 3 shows the comparison of proposed and existing algorithms in terms of recall. The recall of proposed OEIFCM algorithm is 0.97% higher than EIFCM and 2.86% higher than IFCM. From this analysis, it is observed that the proposed OEIFCM algorithm has better recall than the other algorithms.

4.3. F-measure

It is calculated by using both precision and recall as follows:

$$F - measure = 2 \times \left(\frac{precision \times recall}{precision + recall}\right) \quad (20)$$

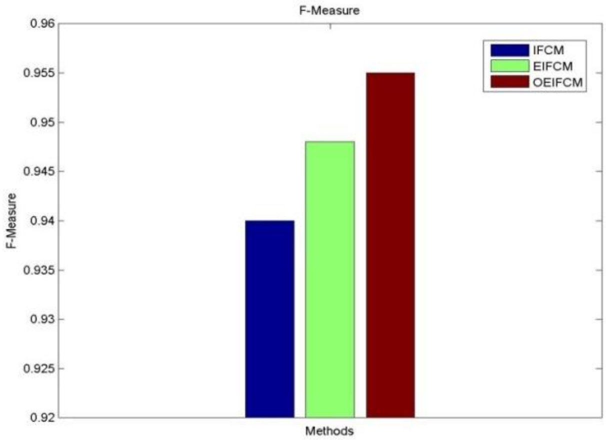


Fig. 4: Comparison of F-Measure.

Figure 4 shows the comparison of proposed and existing algorithms in terms of f-measure. The f-measure of the proposed OEIFCM algorithm is 0.74% higher than EIFCM and 1.59% higher than IFCM. From this analysis, it is observed that the proposed OEIFCM algorithm has better f-measure than the other algorithms.

4.4. Accuracy

It is the fraction of both TP and TN among the total number of cases examined.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{19}$$

Here, TN is the True Negative.

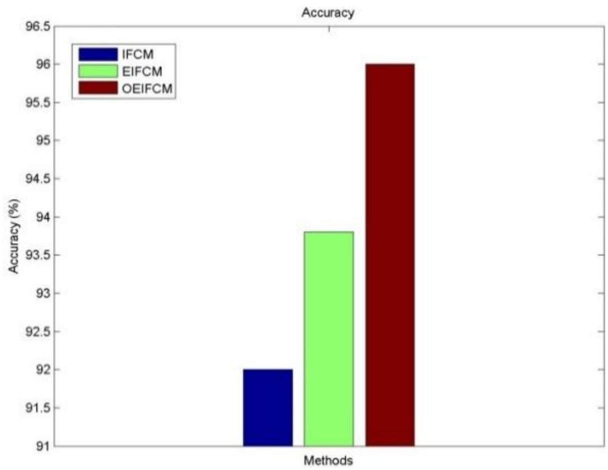


Fig. 5: Comparison of Accuracy.

Figure 5 shows the comparison of proposed and existing algorithms in terms of accuracy. The accuracy of the proposed OEIFCM algorithm is 2.35% higher than EIFCM and 4.35% higher than IFCM. From this analysis, it is observed that the proposed OEIFCM algorithm has high accuracy than the other algorithms.

5. Conclusion

In this article, a performance of groundnut yield prediction is improved by proposing an OEIFCM model. The main aim of this research is to adapt the weight matrix elements based on the learning method and improve the prediction performance more efficiently. In this model, the weight adaptation of IFCM is improved by optimizing the learning parameters based on the SOMA that reduces the computation complexity and improves the learning rate within the minimum number of iterations. Finally, the experimental results are proved that the proposed OEIFCM has better performance than the other crop yield prediction models. The future work involves the selection of the most optimal parameters that improves the prediction accuracy significantly.

References

[1] Papageorgiou EI, Markinos AT, & Gemtos TA (2011), "Fuzzy cognitive map based approach for predicting yield in cotton crop production as a basis for decision support system in precision agriculture application", *Applied Soft Computing*, *11*(4), 3643-3657. https://doi.org/10.1016/j.asoc.2011.01.036.

[2] Ramesh D, & Vardhan BV (2015), "Analysis of crop yield prediction using data mining techniques", *International Journal of Research in Engineering and Technology*, *4*(1), 47-473. https://doi.org/10.15623/ijret.2015.0401071.

[3] Gandge Y (2017), "A study on various data mining techniques for crop yield prediction", in *2017 IEEE International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT)*, pp. 420-423. https://doi.org/10.1109/ICEECCOT.2017.8284541.

[4] Papageorgiou EI, Aggelopoulou KD, Gemtos TA, & Nanos GD (2013), "Yield prediction in apples using Fuzzy Cognitive Map learning approach", *Computers and electronics in agriculture*, *91*, 19-29. https://doi.org/10.1016/j.compag.2012.11.008.

[5] Shastry A, Sanjay HA, & Hegde M (2015), "A parameter based ANFIS model for crop yield prediction", in *2015 IEEE International Advance Computing Conference (IACC)*, pp. 253-257. https://doi.org/10.1109/IADCC.2015.7154708.

[6] Paul M, Vishwakarma SK, & Verma A (2015), "Analysis of soil behaviour and prediction of crop yield using data mining approach", in *2015 IEEE International Conference on Computational Intelligence and Communication Networks (CICN)*, pp. 766-771. https://doi.org/10.1109/CICN.2015.156.

[7] Manjula A, & Narsimha G (2016), "Crop yield prediction with aid of optimal neural network in spatial data mining: new approaches", *International Journal of Information & Computation Technology*, *6*(1), 25-33.

[8] Natarajan R, Subramanian J, & Papageorgiou EI (2016), "Hybrid learning of fuzzy cognitive maps for sugarcane yield classification", *Computers and Electronics in Agriculture*, *127*, 147-157. https://doi.org/10.1016/j.compag.2016.05.016.

[9] Balakrishnan N, & Muthukumarasamy G (2016), "Crop production-ensemble machine learning model for prediction", *International Journal of Computer Science and Software Engineering*, *5*(7), 148.

[10] Silas NM, & Nderu L (2017), "Prediction of tea production in Kenya using clustering and association rule mining techniques", *American Journal of Computer Science and Information Technology*, *5*(2), 1-7. https://doi.org/10.21767/2349-3917.100006.

[11] Garg A, & Garg B (2017), "A robust and novel regression based fuzzy time series algorithm for prediction of rice yield", in *2017 IEEE International Conference on Intelligent Communication and Computational Techniques (ICCT)*, pp. 48-54. https://doi.org/10.1109/INTELCCT.2017.8324019.

[12] Rajak RK, Pawar A, Pendke M, Shinde P, Rathod S, & Devare A (2017), "Crop recommendation system to maximize crop yield using machine learning technique", *International Research Journal of Engineering and Technology*, *4*(12), 950-953.