ENHANCED FUZZY MIN-MAX NEURAL NETWORK WITH PRUNING FOR CLASSIFYING DATA

Jaitra Chakraborti¹, Santosh B. Javheri²

Assoc. Prof², ME student¹, Department of Computer Engineering, JSPM's RSCOE, Pune University, India

ABSTRACT:

Classification is one of the most active research and application area in neural network. Fuzzy Min-Max (FMM) is a neural network model used for classification purpose. There are certain limitations in original FMM neural network regarding formation of hyperboxes. Enhanced Fuzzy Min-Max (EFMM) neural network overcomes the limitations regarding hyperbox expansion, overlap test rule and hyperbox contraction. Due to these new rules EFMM creates more hyperboxes than FMM. Therefore, network complexity of EFMM is more than FMM. EFMM creates large number of hyperboxes which affects the network complexity. So, in this paper, pruning of hyperboxes is done to minimize the hyperboxes and improve the network complexity of EFMM, but the recognition rate decreases slightly. The pruning of hyperboxes is done using confidence factor. The confidence factor of each hyperbox is based on its frequency of use and accuracy of prediction.

Keywords: Classification, Fuzzy min-max model, Hyperbox, Neural Network learning, Pruning

[1] INTRODUCTION

Artificial neural network (ANN) is a computational model that consists of an interconnected group of artificial neurons that replicates the biological neural system in our brain [1]. Pattern classification in current scenario is used for many engineering solutions. Control, tracking, and prediction systems will often use classifiers to determine input-output relationships. Recently, ANNs are used under different fields, e.g., healthcare, power systems and fault detection [2]. Pattern classification is one of the active ANN application domains. For example, ANN models have been successfully applied to classification tasks in business and science and industrial fault detection and diagnosis [3].

In terms of ANN learning, one of the main problem related to batch learning, is catastrophic forgetting [4]. Catastrophic forgetting means the inability of learning system to remember, whenever some new information is given. The catastrophic forgetting is also termed as stability-plasticity dilemma. Stability-plasticity dilemma mainly concerned with how a learning system is plastic enough to learn new information and as well as stable enough to
retain the previously learned information. The issue of stability-plasticity dilemma has to be solved especially when an ANN has to learn from data samples in one pass using online learning.

The FMM neural network is one of the ANN models which overcome the stability-plasticity dilemma. FMM neural network uses hyperboxes for classification. A hyperbox contains patterns with full class membership. A hyperbox is defined by its min point and its max point, and a membership function is defined with respect to these hyperbox min-max points. Learning in FMM classification is characterized by properly placing and adjusting hyperboxes.

Several salient learning properties those are associated with FMM [4].

1) Online learning: The pattern classifier should learn new information without affecting the historical information. Some classifier in neural network uses offline learning in which both the old data and new has to be retrained whenever some new data comes. Online learning is important to solve the stability plasticity dilemma.

2) Nonlinear separability: The pattern classifier should be able to divide two different classes of any shape and size. Non linear separability is the ability to build a nonlinear decision boundary to separate distinct classes. There should be a boundary between patterns of two different classes.

3) Overlapping classes: The ability of the nonlinear decision boundary to reduce misclassification by removing the overlapping regions of different classes. There should not be any overlapping of classified regions of two different classes.

4) Training time: The ability to learn and revise the nonlinear decision boundary with one-pass learning through the training data within a small training time. Some of the algorithms take very large number of passes to learn through the data. The property of pattern classification is to learn the patterns divided by non-linear boundaries in very short time. But when online adaptation is done some problem occurs.

5) Soft and Hard decisions: A pattern classifier must provide soft or hard decisions. A hard decision or crisp decision can be either 0 or 1, i.e. either the pattern belongs to the class or not. A soft decision provides a value that by how much degree the pattern will be of that class.

EFMM [4] has been developed to overcome the limitations of FMM. Three heuristic rules in EFMM to overcome the limitations are:
1. To eliminate the problem of overlapping during hyperbox expansion, new overlapping rules has been suggested.
2. To discover other overlapping cases the hyperbox test rule has been extended.
3. To resolve the hyperbox overlapping cases, hyperbox contraction rule is provided.

But it has been observed that the network complexity of EFMM is more than FMM due to formation of large number of hyperboxes. So, to minimize the network complexity of EFMM the concept of pruning is introduced in this paper. We do pruning of hyperboxes to remove the unnecessary hyperboxes containing less accuracy or having less confidence factor related to a class. The confidence factor will depend on the recent uses of the hyperboxes and the prediction accuracy of the hyperboxes.

[2] LITERATURE SURVEY

The traditional systems describes crisp events, i.e. either it will occur or don’t. Fuzzy set was introduced by Zadeh [5] to represent and manipulate data that are not accurate, rather fuzzy. After the fuzzy set has been introduced it has been used for solving problems related to classification and clustering.

For higher level of decision making a fuzzy set approach for classification of pattern is used. First, FMM (Classification) network [6] proposed by P.K. Simpson, is used to serve for supervised learning. In this [6] relationship between the pattern classification and fuzzy sets has been described. It explains how the fuzzy min-max classifier neural network works using learning and recalling algorithm. A neural network classifier that uses min-max hyperboxes as fuzzy sets that are combined into fuzzy set classes was introduced.

P.K. Simpson also introduces FMM (Clustering) network [7] that is used to serve for unsupervised learning. There were several faults in the original fuzzy ART that has been corrected in this.

General Fuzzy Min-Max (GFMM) algorithm [8] which is the fusion of both classification and clustering related to FMM has been introduced by Gabris and Bargiela. It is developed on the basis of expansion and contraction process. This algorithm can be applied as pure classification, pure clustering or hybrid clustering classification. In this [8] the classification results can be crisp or fuzzy. Similarly to the original methods, the GFMM utilizes min-max hyperboxes as fuzzy sets. In GFMM, the training of data is very fast, and as long as no identical data belongs to different classes, the recognition rate is 100%.

In Stochastic FMM [9] proposed by A. Likas, the fuzzy min-max neural network [5] can be trained increasingly by appropriately adjusting the number of hyperboxes and their corresponding values. In this [9] the idea of random hyperbox and stochastic fuzzy min-max neural network has been proposed. Each hyperboxes is associated with stochastic learning automaton.

Rizzi et al. introduces Adaptive resolution min-max classifier [10], in which two algorithms Adaptive Resolution Classifier (ARC) and pruned ARC is devised. The automation degree of the training procedure is also an important factor with generalization capability and noise robustness. The generalization capability of the original min-max classifier depends mostly on position and size of the hyperboxes generated during training. In this [10] the
classification system is automatic, since the training algorithm does not depend on presentation order of pattern and no critical parameter must be set by the user.

Inclusion/Exclusion fuzzy hyperbox classifier [11] is described in which, one or more fuzzy hyperbox defined by their corresponding minimum and maximum vertices and the hyperbox membership function is used to describe each class. Inclusion hyperbox and exclusion hyperbox are the two types of hyperboxes created. With these two types of hyperboxes each class fuzzy set is represented as a combination of inclusion hyperboxes of the same class minus a combination of exclusion hyperboxes.

A new model called the Fuzzy Min-Max Neural Network Classifier with Compensatory Neurons (FMCN) was proposed. To represent the pattern classes, FMM with Compensatory Neuron (FMCN) hyperbox fuzzy sets was used. The concept of compensatory neuron [12] comes from how human brain works at difficult conditions. The use of the contraction process is avoided by FCMN, reduces errors caused by it. Since the hyperboxes that are already created are not contracted, FMCN can retain the knowledge of the already learned patterns more efficiently than FMNN and GFMN. It is [12] based on the reflex mechanism of human brain, learns the data online in a single pass of data, and maintains simplicity.

A data-core-based FMM neural network (DCFHN) model for pattern classification was proposed [13]. For classifying the neurons of DCFMN, a new membership function has been defined in which the data core, the noise and the geometric center of the hyperbox are considered. A new membership function [13] based on the data core is used instead of contraction process of the FMNN. To show the overlapping areas of hyperboxes of different classes, the membership function is added to neural network.

Quteishat and Lim proposed new algorithm named Modified FMM. In attempt to improve the classification performance of FMM, few numbers of large hyperboxes are formed in the Modified FMM (MFMM) network [14] and some modifications are done. A new input pattern is given; Euclidean distance measure is used for predicting the target class associated with the new input and also the fuzzy membership function of the input pattern to the hyperboxes formed in FMM has to be measured. In this [14] a rule extraction algorithm is also enclosed. For each FMM hyperbox a confidence factor is calculated, and a user-defined threshold is used to prune the hyperboxes with low confidence factors.

Modified FMM with Genetic Algorithm (MFMM-GA) [15] is a two stage classification of pattern and extraction of rule process. The first stage consists of Modified Fuzzy min-max (MFMM) classifier and the second stage is based on the genetic algorithm (GA) based classifier. To reduce the number of features in the extracted rules, a “don’t care” approach is selected by the GA rule extractor and fuzzy if–then rules are extracted from the modified FMM classifier.

Offline and Online FMM and the classification and regression tree (CART), a new approach [16, 17] to classify and detect faults using a hybrid fuzzy min-max (FMM) neural network and classification and regression tree has been proposed. It [16] uses the concept of FMM for the purpose of classification and CART is used for rule extraction process. It also supports the offline and online learning properties for fault detection and diagnosis process.
[3] EFMM LEARNING ALGORITHM

The EFMM [4] consists of three steps i.e. hyperbox expansion, overlap test rule and hyperbox contraction. Learning in FMM begins with datasets consisting of input patterns and target classes. FMM creates hyperboxes based on the input patterns. Hyperboxes are represented by minimum and maximum points in an n-dimensional space within a unit hypercube ($I^n$).

Hyperbox fuzzy sets are defined as

$$B_j = \{(A_{h}, v_j, w_j, f((A_{h}, v_j, w_j)))\}/\forall A_h \in I^n$$  \hspace{1cm} (1)

where, $B_j$ is the hyperbox fuzzy set, $A_h = (a_{h1}, a_{h2}, \ldots, a_{hn})$ is the input pattern, and $v_j = (v_{j1}, v_{j2}, \ldots, v_{jn})$ and $w_j = (w_{j1}, w_{j2}, \ldots, w_{jn})$ are the minimum and maximum points of $B_j$, respectively.

The hyperboxes will be formed as per the input given to the system. Formation of hyperbox for new input patterns is depend upon the membership function, given by

$$B_j(A_h) = \frac{1}{2n} \sum_{i=1}^{n} \left[ \max(0, 1 - \max(0, \min(1, a_{hi} - w_{ji})) + \max(0, 1 - \max(0, \min(1, v_{ji} - a_{hi}))) \right]$$  \hspace{1cm} (2)

where $B_j$ is the membership function of the $j$th hyperbox, $A_h = (a_{h1}, a_{h2}, \ldots, a_{hn}) \in I^n$ is the $h_{th}$ input pattern, and $y$ is a sensitivity parameter that regulates how fast the membership decreases as the distance between $A_h$ and $B_j$ increases.

After the hyperboxes are generated, then which hyperbox will belong to which class is decided by matrix $U$, as

$$u_{jk} = \begin{cases}  \text{1 if } b_j \text{ is hyperbox for class } C_k \\ \text{0 otherwise} \end{cases}$$ \hspace{1cm} (3)

where $b_j$ is the $j$th hyperbox node and $C_k$ is the $k$th class node.

This can be represented as the following figure.

![Figure 2: Three-layer FMM network](image)

$C_k$ can be calculated as
ENHANCED FUZZY MIN-MAX NEURAL NETWORK WITH PRUNING FOR CLASSIFYING DATA

\[ C_k = \max_{1 \leq j \leq m} b_j u_{jk} \]  

(4)

for each class nodes.

The FMM structure consists of three layers. \( F_A \) is the input layer, which contains the input patterns from the user. \( F_B \) is the hyperbox layer, which shows the hyperbox created during learning process. The connection between \( F_A \) and \( F_B \) is shown by the minimum (\( V \)) and maximum (\( W \)) points of the hyperbox. \( F_C \) is the output layer which represents the class. The connection between \( F_B \) and \( F_C \) are binary values and are stored in matrix \( U \).

In EFMM, the input dataset is provided by the user. During learning phase, hyperboxes are created or assigned to the input dataset. The hyperbox expansion rule is provided, which uses expansion coefficient (\( \Theta \)) for expansion.

\[ \text{Max}_n(W_{ji}, a_{hi}) - \text{Min}_n(V_{ji}, a_{hi}) \leq \Theta \]  

(5)

Based on the above equation it is checked that how much a hyperbox can be expanded with respect to expansion coefficient (\( \Theta \)). In FMM, during expansion process it computes the sum of all dimensions and compares with \( (n\Theta) \). Whereas, EFMM considers each dimension individually and checks the difference between minimum and maximum points of each dimension separately against \( \Theta \).

![Figure 3: 3-D Hyperbox](image)

Hyperbox Overlap Test Rule: In FMM, the cases were not sufficient to identify all the overlapping cases. So, new cases are added in EFMM to overcome this problem. Overlapping exists when one of the following cases given in [4] is satisfied. Due to the new overlapping rules in EFMM, the misclassification reduces and also causes generation of more number of hyperboxes.

Hyperbox Contraction Rule: The contraction rule is given based on the cases of the hyperbox overlap test. Here, all cases are tested to find a proper adjustment.

Network Pruning: After the EFMM network is trained, pruning is used to reduce the number of hyperboxes. As the number of hyperboxes gets reduced, the rule extraction process will be faster. The pruning of hyperboxes is done using confidence factor. The confidence factor of each hyperbox is based on its frequency of use and accuracy of prediction.

\[ CF_j = (1 - w)F_j + wA_j \]  

(6)
where, $F_j$ is the frequency of use of hyperbox $j$, $A_j$ the accuracy of prediction of hyperbox $j$, and $w \in [0, 1]$ is a weighing factor.

The value of $F_j$ is defined as the number of patterns in the prediction set classified by any hyperbox $j$, divided by the maximum number of patterns in the prediction set classified by any hyperbox with the same classification class.

On the other hand, the value of $A_j$ is defined as the number of correctly predicted set of patterns classified by any hyperbox $j$, divided by the maximum correctly classified patterns with the same classification class.

The confidence factor finds the good hyperboxes that are frequently used and gives high classification accuracy, and also the hyperboxes that are rarely used but still gives the high classification accuracy.

[4] EXPERIMENTAL RESULTS

The dataset used in this paper is the IRIS dataset. IRIS dataset is available freely and can be accessed from UCI machine learning repository. The IRIS dataset has 4 attributes and 3 classes. There are 150 instances available to use. Each class contains 50 instances. This complete dataset is divided into training dataset, prediction dataset and testing dataset for classification.

We have implemented the EFMM neural network for learning and classification of data. To test the performance of FMM with EFMM and EFMM with pruning we have used Iris data set from UCI machine learning repository. For training of neural network 60% data is used, 20% data is used as the prediction set. Then for testing complete dataset is used.

To check the total number of hyperboxes formed, the value of expansion parameter is varied from 0.05 to 0.95. The sensitivity parameter for the membership function is set to 0.5. For pruning of hyperboxes the cut of value is set to 0.5 and weighing factor is set to 1.

Table 1 shows the total number of hyperboxes obtained when $\Theta$ is to 0.3, 0.4, 0.5 and 0.6.

<table>
<thead>
<tr>
<th>$\Theta$</th>
<th>Number of Hyperboxes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FMM</td>
</tr>
<tr>
<td>0.3</td>
<td>39</td>
</tr>
<tr>
<td>0.4</td>
<td>28</td>
</tr>
<tr>
<td>0.5</td>
<td>20</td>
</tr>
<tr>
<td>0.6</td>
<td>16</td>
</tr>
</tbody>
</table>

Figure 4 shows the graph generated for FMM, EFMM and EFMM after pruning for different values of $\Theta$. 
ENHANCED FUZZY MIN-MAX NEURAL NETWORK WITH PRUNING FOR CLASSIFYING DATA

It has been found that the number of hyperboxes created in EFMM is more than that of FMM. This is because of the new rules introduced in EFMM [4]. But EFMM has less misclassification rate than that of FMM. Figure 5 shows the recognition rate of FMM, EFMM and EFMM after pruning for Iris data set. Table 2 shows the average of recognition rate.

Table 2: Average of recognition rate

<table>
<thead>
<tr>
<th>Dataset</th>
<th>FMM</th>
<th>EFMM</th>
<th>EFMM after pruning</th>
</tr>
</thead>
<tbody>
<tr>
<td>IRIS</td>
<td>98.9 ± 1.1</td>
<td>99.2 ± 1.7</td>
<td>97.7 ± 2.15</td>
</tr>
</tbody>
</table>
[7] CONCLUSION

EFMM neural network is used for classification purpose. EFMM uses the concept of hyperbox for classification. There are four steps followed in EFMM neural network with pruning. First, the hyperbox is expanded to contain the inputs. Second, overlap of hyperboxes is checked. Third, Contraction is done according to the overlaps. At last, EFMM creates more number of hyperboxes as compared to FMM neural network. As the number of hyperboxes is more, the network complexity of EFMM is more. So, in this project, pruning of hyperboxes is done over EFMM neural network. Pruning of hyperboxes is done using the concept of confidence factor. The confidence factor of each hyperbox is based on its frequency of use and accuracy of prediction. As the number of hyperboxes reduced the network complexity of EFMM also reduces, but it affects the recognition ratio by small margin.
REFERENCES