Fuzzy Hyper-line Segment Neural Network by using Association Rule Mining

B. S. Shetty1*, U. V. Kulkarni2, Preetee M. Sonule3, Manisha N. Shinde4

1Department of Information Technology- SGGSIET- SRTMU Nanded, MH, India
2 Department of Computer Engineering- SGGSIET -SRTMU Nanded, MH, India
3 Department of Information Technology- SGGSIET -SRTMU Nanded, MH, India
4Department of Information Technology- SGGSIET -SRTMU Nanded, MH, India

*Corresponding Author: bsshetty@sggs.ac.in, Tel.:7276300495
Available online at: www.ijcseonline.org
Accepted: 23/Dec/2018, Published: 31/Dec/2018

Abstract— In this paper, we have proposed the fuzzy hyper-line segment neural network (FHLSSNN) by using association rule mining(FHLARM). Regression tree is used for pattern recognition. We have used supervised learning neural network classifier for classification of fuzzy sets. The FHLARM make the pattern classification with the help of hyper-line segments. It has two endpoints and corresponding member-ship function. The proposed model is evaluated by using iris, wine and solar mine datasets. For extraction of rules, we have used association rule mining. It gives the better classification accuracy results on various datasets as compared to previous methods. Regression tree maintains a hierarchy of extracting rules.

Keywords—Fuzzy sets, Neural Network, Supervised and unsupervised methods, Pattern classification, FMM.

I. INTRODUCTION

A computer system which is based on the human brain and nervous system is called Neural Network. The membership function is used to determine the nature of its neurons by using learning rule and by the architecture itself. In detection of pattern use the fuzzy min-max neural networks (FMN) are widely used and they become famous. The supervised Learning and Unsupervised Learning are the two main types of FMN. In supervised method, the classes have some labels and they having input patterns and end points which helps to reduce the misclassification [2]. While in unsupervised learning, there is no any label to patterns but it forms some groups called clusters has suitable data. Supervised and unsupervised learning are also called as classification and clustering methods respectively. In super-vised learning we have some apriori knowledge of object but in unsupervised learning we do not have any apriori knowledge [1]. Learning in FMM starts using a dataset consisting of input patterns and uses membership functions. A curve having membership value between 0 to 1 is called membership function. The various applications of neural networks are speech detection, pattern recognition, signature verification, detection of human face etc [2]. Our motivation is to extract the rules using this innovative approach.

The remaining portion of paper is explained as follows, in section II, we have given the brief idea about Related Works, in section III, Learning algorithm of FHLSSNN is given. Section IV, the proposed modified approach shows the association rule mining for extracting rules. In section V, regression tree is explained for FARM, Results and comparison is explained in section VI, Section VII shows one 2-dimensional example for FARM. Section VIII shows conclusion.

II. RELATED WORK

The FHLSSNN architecture consists of four layers $F_R, F_E, F_D$ and $F_C$ respectively as shown in Figure 1. The $F_R$ layer is the lowest layer as shown in figure, which consists of $n$ processing elements, accepts an input pattern. At the time of training $F_R$ shows $m$ processing nodes. The connections from $F_R$ to $F_E$ shows two end points of line, Figure 2 shows the implementation of fuzzy hyper-line segment.

$F_R$ is a layer which accepts input pattern.

$F_E$ is a layer which having $m$ processing nodes.

$F_D$ is a layer which shows membership function.

$F_C$ is a layer which shows number of classes.

Let $R_h = (r_{h1}, r_{h2}, ...., r_{hn})$ is the input node.

The end points of $V_j$ which are $V_j = (v_{j1}, v_{j2}, ...., v_{jn})$ and $W_j = (w_{j1}, w_{j2}, ...., w_{jn})$

The membership function of $j^{th}$ $F_E$ node is given below:

eq.1
\[ e_j(R_h, V_j, W_j) = 1 - f(x, y, l) \]  

Here, \( x = l_1 + l_2 \) and distances \( l_1, l_2 \) and \( l \) are computed by using following formulae eq.2, eq.3, eq.4

\[ l_1 = \left( \sum_{i=1}^{n} (w_{ji} - r_{h1})^2 \right)^{1/2} \]  

\[ l_2 = \left( \sum_{i=1}^{n} (v_{ji} - r_{h2})^2 \right)^{1/2} \]  

\[ l = \left( \sum_{i=1}^{n} (v_{ji} - r_{hi})^2 \right)^{1/2} \]

Where, \( l_1 \) is length of \( R_h \) from \( w_{ji} \), \( l_2 \) is length of \( R_h \) from \( v_{ji} \), and \( l \) is length of hyperline segment.

Since, \( f() \) is a three-parameter ramp threshold function defined as

\[ f(x, y, l) = \begin{cases} xy & \text{if } 0 \leq xy \leq 1 \\ 1 & \text{if } xy > 1 \end{cases} \]

Every node of \( F_C \) and \( F_D \) layer shows a class. The \( F_D \) layer gives choice. The result of \( k^{th} \) \( F_D \) node shows how much the pattern has a place with the class \( d_k \).

### III. FHLSNN LEARNING ALGORITHM METHODOLOGY

This algorithm is used to create fuzzy hyper-line segment which is having three stages, First creates the hyper-line segment, the second checks their intersection and third remove intersections. If two lines have intersected, they that are not overlapped [7].

#### A. Creation of hyper-line segment

The length of hyper-line section is limited by parameter \( \xi \), \( 0 \leq \xi \leq \xi_{max} \) and \( \xi_{max} \) relies upon the measurement of highlight vector. In learning procedure suitable value of \( \xi \) is chosen and hype-line is expanded when length of hyper-line portion after expansion is not equivalent to \( \xi \). Expecting training set described as \( R \in \{ R_h | h = 1, 2, \ldots, p \} \). Learning begins by applying inputs one by one from pattern set \( R \). Given training pair \( (R_h, d_h) \), search all hyper-line segments of Class \( d_h \). The following four sub-stages are carried one by one for possible consideration of input pattern \( R_h \).

#### Sub-stage 1: Check whether pattern falls on any of the hyper-line sections. Which can be confirmed by utilizing fuzzy hyper-line portion membership work as given in eq.(1).

For \( k = 1, 2, \ldots, p \) and \( j = 1, 2, \ldots, m \)

\[ u_{jk} = \begin{cases} 1 & \text{if } e_j \text{ is a segment of hyper-line of class } d_k \\ 0 & \text{otherwise} \end{cases} \]

Where \( e_j \) shows \( j^{th} \) \( F_E \) node and \( d_k \) shows \( k^{th} \) \( F_D \) node.

The membership function of every \( F_D \) node union performs the union with fuzzy values.

\[ d_k = \max_{j=1}^{m} u_{jk} \text{ for } k = 1, 2, \ldots, p \]

Every \( F_C \) node delivers non-fuzzy output described as in eq.(7)

\[ c_k = \begin{cases} 0 & \text{if } d_k < T \\ 1 & \text{if } d_k = T \end{cases} \]

Where \( T = \max (d_k) \), for \( k = 1 \) to \( p \).
Sub-stage 2: If the pattern falls on any of hyper-line then that FHLS is extended to include that pattern. Consider $e_j$ is that hyper-line segment with end points $V_j$ and $W_j$ then $l_1$, $l_2$ and $l$ are computed using equations (2), (3) and (4).

2(a): Check whether $l_2 > l_1$ and test the point $v_j$ falls on hyper-line segment formed by points $W_j$ and $R_h$. If $e_j(W_j, R_h, W_j) = 1$, then hyper-line segment is extended by changing end point $v_j$ by $R_h$. To include $R_h$ if extension criteria is satisfied then the new points are

\[ V_j^{new} = R_h \ and \ W_j^{new} = W_j \]

2(b): Check whether $l_2 > l_1$, and test whether the point $V_j$ falls on hyper-line segment formed by points $v_j$ and $R_h$. If $e_j(W_j, V_j, R_h) = 1$, then hyper-line segment is extended by changing end point with $R_h$, $W_j$ to include $R_h$, if extension criteria is satisfied, hence eq. shows

\[ W_j^{new} = R_h \ and \ V_j^{new} = V_j \]

Sub-stage 3: To include pattern $R_h$ we have to expand available point which is considered as a hyper-line. That is expanded as

\[ V_j^{new} = R_h \ and \ W_j^{new} = W_j \]

Sub-stage 4: If pattern $R_h$ is not included by any of above sub-stages then new hyper-line segment is created for that class which is described as

\[ W_{new} = R_h \ and \ V_{new} = R_h \]

B. Intersection Test

The intersection of hyper-lines has been supported by learning algorithm having same class. For different classes intersection is eliminated and the hyper-line is extended. Let, $W_{1st} = [x_1, x_2, ... x_n]$, and $V_{1st} = [y_1, y_2, ... y_n]$ represent two end points of extended hyper-line segment of one class and $W_n = [x_1, x_2, ... x_n]$ and $V_n = [y_1, y_2, ... y_n]$ are end points of hyper-line segment of another class. Check whether the hyper-line passing through two segments is intersected or not. This is described by following equations. The equation of hyper-line passing through $W_{1st}$ and $V_{1st}$ is

\[ \frac{x_i - x_j}{y_i - y_j} = r_1 \ for \ i = 1, 2, ..., n \]

Equation of hyper-line passing through $W_n$ and $V_n$ is

\[ \frac{x_i - x_j}{y_i - y_j} = r_2 \ for \ i = 1, 2, ..., n \]

Where $r_1$ and $r_2$ are the constants where $a_i$ and $b_i$ are the variables.

The equations (5) and (6) leads to set of $n$ simultaneous equations which are described as

\[ r_1(y_i - x_i) + x_i = r_2(y_i - x_i) + x_i \]

For $i = 1, 2, ..., n$

We have to use two simultaneous equations to find $r_1$ and $r_2$. If remaining equations are satisfied by using values of $r_1$ and $r_2$ then two hyperlines has intersected with point of intersection $P_t$ which is calculated by eq.8.

\[ p_t = (r_1(y_1 - x_1) + x_1, ..., r_1(y_n - x_n) + x_n) \]

C. Removing Overlap

The segments which are from dissimilar classes are intersected in sub-stage 2(a) and sub-stage 3. This intersection can be removed by restoring points $V_j$ as $V_j^{new} = V_j$. And $W_j$ as $W_j^{new} = W_j$ is used to remove intersection of end points. Those points are formed by sub-stage 2(b). After that a new hyper-line segment is created which include input pattern $R_h$. In sub-stage 4 the intersection is removed by restoring end points which are $W_{new + 1} = W_{new + 1}$ and $V_n = W_n$

IV. RULE EXTRACTION FOR FHLS ARM

Association rule mining (ARM) is a vital research point in information mining field of learning explore. The information aggregation, huge numbers of business are increasingly designed for mining association rules from their databases. Association control mining calculation is primary substance of association govern mining research. Apriori algorithm is one of the best calculations for association rule mining. Apriori method needs to filter the first database for commonly in the mining procedure. Also it creates an extensive number of competitor thing sets. There are low mining effectiveness, too vast possessed memory space and different insufficiencies, needs facilitate study [21].

Apriori algorithm is the most generally utilized method in association rules. Association rules mining is expressed while breaking down the retail exchange databases, the present advancement has incredibly surpassed the first extent of use, the depth and breadth have been significantly upgraded [24]. The fundamental model of association rule mining is appeared in Fig3.

In FARM, Association mining is a strategy which is intended to discover visit designs, connections, associations, or causal structures from informational indexes. They are found in number of databases, for example, social databases, value-based databases. An arrangement of exchanges is given, association rule mining intends to discover the
guidelines which empower us to foresee the event of a particular thing in light of the events of other things in the transactions [21].

A. Apriori Algorithm

In FHLARM, to reduce the quantity of applicant information thing sets, Apriori calculation utilizes layer look strategy. Also it examines all the exchange information things of the database D to locate the regular thing sets.

From the highlights of successive thing sets it can be realized that only the superset created by those hopeful sets which have been confirmed as frequent items might be visit thing set

In this way, just to utilize frequent item sets to create the competitor set for next sweep, that is, utilizing Li-1 to produce Ci, in which, L is the arrangement of successive thing sets, C is the arrangement of hopeful things. Every database just thinks about the set with a similar number of information. Apriori algorithm can create a moderately little hopeful information thing sets. The competitor information thing set does not need to be over and again produced by the records in the database. They are produced by k-1 visit information thing set which is created by the last flow during the time spent. Which is looking for k visit information thing sets. The Apriori algorithm execution process as shown in Figure.4 [21].

V. REGRESSION TREE FOR FHLARM

The FHLARM having a tree structure which is used to display characterization that is called regression tree (RT). It divides a dataset into little subsets with increment inside and out of tree. The final conclusion is a tree with choices and leaves. A choice node (for example, Outlook) has at least two branches (for example, Sunny, Overcast and Rainy). Leaf node (for example, Play) speaks to a classification or choice. The root node in the tree having greatest value. Choice trees can deal with both absolute and numerical information. By using weka 3.6 we have plot trees for extracted rules. By using ARFF format of Iris dataset we can draw any type of tree. For example J48 regression tree using weka 3.6 is shown below.

J48 regression tree
------------------
Petalwidth <= 0.0641: 1 (50.0)
Petalwidth > 0.0641
| Petalwidth <= 0.2051
| | Petalwidth <= 0.8718
| | | petalength <= 0.6154: 2 (47.0/1.0)
| | | petalength > 0.6154
| | | | petalength <= 0.1795: 3 (3.0)
| | | | petalength > 0.1795: 2 (2.0)
| | Sepallength > 0.8718: 1 (2.0/1.0)
Petalwidth > 0.2051: 3 (46.0/2.0)

Here the size of tree is 11 and number of leaves are 6. As like iris we can find this tree for different datasets.

Finally, all results and comparisons for different classifiers with different datasets are shown in below table.
VI. RESULTS AND COMPARISION

The pattern recognition of FHLSNN algorithm are compared with fuzzy min-max neural network. Finally, we got the FARM gives best pattern recognition on various datasets by comparing different algorithms. In the below table, the performance of FARM algorithm is compared with different classifiers for the training and recall time per pattern. The association rule and regression tree forms rules and tree. To check the accuracy on different datasets and comparison between them check the below tables.

Table 1: Iris dataset

<table>
<thead>
<tr>
<th>Train Data (%)</th>
<th>100-100</th>
<th>80-20</th>
<th>70-30</th>
<th>50-50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
<td>FHLARM</td>
<td>FHLARM</td>
<td>FHLARM</td>
<td>FHLARM</td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>99.1</td>
<td>98.2</td>
<td>98.8</td>
<td>99.2</td>
</tr>
<tr>
<td>No. of rules</td>
<td>8</td>
<td>7</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Instances attributes</td>
<td>150-4</td>
<td>150-4</td>
<td>150-4</td>
<td>150-4</td>
</tr>
</tbody>
</table>

Table 2: Solar mine dataset

<table>
<thead>
<tr>
<th>Train Data (%)</th>
<th>100-100</th>
<th>80-20</th>
<th>70-30</th>
<th>50-50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
<td>FHLARM</td>
<td>FHLARM</td>
<td>FHLARM</td>
<td>FHLARM</td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>97.7</td>
<td>98.7</td>
<td>97.9</td>
<td>99.4</td>
</tr>
<tr>
<td>No. of rules</td>
<td>12</td>
<td>10</td>
<td>11</td>
<td>8</td>
</tr>
<tr>
<td>Instances attributes</td>
<td>300-4</td>
<td>300-4</td>
<td>300-4</td>
<td>300-4</td>
</tr>
</tbody>
</table>

Table 3: Wine dataset

<table>
<thead>
<tr>
<th>Train Data (%)</th>
<th>100-100</th>
<th>80-20</th>
<th>70-30</th>
<th>50-50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method</td>
<td>FHLARM</td>
<td>FHLARM</td>
<td>FHLARM</td>
<td>FHLARM</td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>99</td>
<td>98.7</td>
<td>98.7</td>
<td>97.8</td>
</tr>
<tr>
<td>No. of rules</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Instances attributes</td>
<td>177-13</td>
<td>177-13</td>
<td>177-13</td>
<td>177-13</td>
</tr>
</tbody>
</table>

Table 4: Comparison of classifiers with other classifiers

<table>
<thead>
<tr>
<th>Methods</th>
<th>Misclassification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bayes Classifier</td>
<td>2</td>
</tr>
<tr>
<td>K nearest neighbour</td>
<td>4</td>
</tr>
<tr>
<td>Fuzzy k-NN</td>
<td>4</td>
</tr>
<tr>
<td>Perceptron</td>
<td>3</td>
</tr>
<tr>
<td>Fuzzy Min-Max NN</td>
<td>2</td>
</tr>
<tr>
<td>GFMN</td>
<td>1/0</td>
</tr>
<tr>
<td>M-FMMN</td>
<td>1/0</td>
</tr>
<tr>
<td>EFNN</td>
<td>0</td>
</tr>
<tr>
<td>FARM</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 5: 2D Dataset [14]

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Input pattern(Rh)</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0, 0</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>0.15, 0.05</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>0.2, 0.15</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>0.3, 0.15</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>0.5, 0.2</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>0.3, 0.25</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>0.15, 0.3</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>0.4, 0.3</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>0.4, 0.4</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>0.4, 0.5</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>0.6, 0.5</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>0.6, 0.6</td>
<td>2</td>
</tr>
</tbody>
</table>

VII. 2D SPACE EXAMPLE FOR TWO CLASSES

To describe the nature of FHLARMT we have taken a 2-dimensional example which having 12 patterns shown in below table. Patterns of same class are not close to each other.

Below Figure 5 shows the distribution of patterns.

These twelve patterns gives number of hyperline segments. The below Figure 6 shows the hyperline segments which are formed by removing the intersection of given patterns.
The HLS in above figure are formed by using FHLSNN learning algorithm.
By using FARM algorithm, we can find number of rules for these patterns which are as follows:

Rule (Support, Confidence)
Student1 -> Student3 (50%, Inf%)
Student3 -> Student1 (50%, 100%)
Student2 -> Student3 (50%, Inf%)
Student3 -> Student2 (50%, 100%)

For this demo we have used Zachary's karate network of friendships among 34 students of the club [25].
In this context an association rule Student1, Student2 -> Student3 will mean that students that are friends with Student1 and Student2 are very likely to be friends with Student3 as well, therefore can be used to find missing links in the network.
Although the karate network is undirected (if a is connected to be then be is connected to a), the direction of the resulting rules are important (a -> b means students who are friends with a are also friends with b, but not the other way around).
By using iris dataset there are 8 number of rules are formed for same example which are as follow
Rule (Support, Confidence)
Petalwidth-> Petalength (66%, 9900%)
Petalength-> Petalwidth (66%, 194.1176%)
Sepallength-> Petalength (66%, Inf%)
Petalength-> Sepallength (66%, 194.1176%)
Sepallength-> Petalength (66%, 194.1176%)
Petalength-> Sepallength (66%, 194.1176%)
Sepalwidth-> Petalength (66%, Inf%)
Petalength-> Sepalwidth (66%, 194.1176%)

Table 6: Comparison of ARM with different Algorithms:

<table>
<thead>
<tr>
<th>Methods</th>
<th>Iris</th>
<th>Wine</th>
<th>Solar mine</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIS</td>
<td>89.38</td>
<td>91.22</td>
<td>92.05</td>
</tr>
<tr>
<td>FP-growth</td>
<td>90.07</td>
<td>91.89</td>
<td>92.97</td>
</tr>
<tr>
<td>Apriori</td>
<td>92.87</td>
<td>94.03</td>
<td>93.18</td>
</tr>
<tr>
<td>ARM</td>
<td>94.00</td>
<td>95.33</td>
<td>94.81</td>
</tr>
</tbody>
</table>

The above table shows the comparison of classification accuracy on different datasets using different algorithms. The AIS algorithm is proposed by Agrawal, Imielinski, and Swami for mining association rule. This algorithm is used to improve the quality of dataset together with their functionality. The FP-growth algorithm is used to mining the frequent itemsets and convert them into constructing tree. The apriori algorithm is used to mine frequent itemsets from transactional database[26].

VIII. CONCLUSION
In this paper, we have proposed that the innovative FHLARM learning algorithm. This method is improved by using association rule mining method. The ARM is used to extracting rule and find out different results. The Regression tree shows the hierarchy of extracted rules.

REFERENCES


Authors Profile

Mr. Balaji S Shetty has completed ME in Computer Engineering from Swami Ramanand Teerth Marathwada University Nanded, Maharashtra, India. He is currently working as Asst. Professor in IT Department at SGGSIE&T (Autonomous), Nanded, Maharashtra, India. He has more than 15 years of teaching experience.

Dr. U. V. Kulkarni completed Ph.D. in Electronics and Computer Science Engineering in 2002 from Swami Ramanand Teerth Marathwada University Nanded, Maharashtra, India. He works as Professor and Head in Computer Science and Engineering Department at SGGSIE&T (Autonomous), Nanded, Maharashtra, India. His teaching experience is more than 30.

Ms. Preetee M. Sonule has completed her B. Tech in Computer and M. Tech in IT in SGGSIE&T (Autonomous), Nanded, Maharashtra, India. College comes under Swami Ramanand Teerth Marathwada University Nanded, Maharashtra, India. She has two teaching experience.

Ms. Manisha N. Shinde has completed her M. Tech in IT in SGGSIE&T (Autonomous), Nanded, Maharashtra, India. College comes under Swami Ramanand Teerth Marathwada University Nanded, Maharashtra, India.