ISSN XXXX XXXX © 2017 IJESC

Research Article Volume 7 Issue No.5

Neuro-Fuzzy Classifier System Based on Class-Based Grouping and EM Clustering Algorithm

Heena .K. J. Shaikh[1], D. D. Dharmadhikari[2], S. A. Kinariwala[3]
Department of Computer Science and Engineering
MIT College, Aurangabad, India

Abstract:
A Neuro Fuzzy System is a Fuzzy System that uses learning algorithm borrowed from or influenced by Neural Network Theory to resolve its parameters (fuzzy sets and fuzzy rules) by transform data samples. Existing system focused on Neuro Fuzzy System Classification using Class-Based Grouping Algorithm. In the Class-Based Grouping Process, the parameter values like centroid, threshold value and standard deviation are estimated. These parameters are used for creating the clusters. Using this method a large number of clusters are generated. In proposed work Expectation Maximization (EM) Clustering is used to reduce cluster size. EM Clustering Algorithm for calculating exact mean and standard deviation. After that Transition and Classification process is performed. The EM Clustering gives the improved accuracy and also cluster size of each feature is reduced as compared to Class-Based Algorithm. Finally Accuracy, Mean Absolute Error these parameters are used to evaluate performance of Class-Based Grouping Algorithm and EM Clustering Algorithm. The Class-Based Grouping Algorithm and EM Clustering Algorithm are implemented on Abalone, Diabetes dataset. Accuracy Class-Based Grouping Algorithm is 70.00 and EM algorithm is 76.61 using Abalone dataset. Accuracy of Class-Based Grouping Algorithm is 91.32 and EM Algorithm is 95.10 using Diabetes dataset.

Keywords: EM Clustering Algorithm, Class-Based Grouping Algorithm, Neural Network.

I. INTRODUCTION

Artificial Intelligence (AI) is intelligence conferred by machines. In computer science, the field of AI exploration describe itself as the survey of "intelligent agents" any device that characterize its environment and takes actions that maximize its chance of success at some goal [1]. The term "artificial intelligence" is prescribed when a machine mime "cognitive" functions that humans accomplice with other human minds, such as "learning" and "problem solving" (known as Machine Learning) [2-3] Artificial Neural Networks (ANNs) often just called a Neural Network is mathematical model model based on Biological Neural Network. Artificial Neural Networks are a computational model recycled in computer science and other research disciplines, which is based on a large collection of simple neural units (artificial neurons), loosely comparable to the recognized behavior of a biological brain's axons [4]. Each neural unit is correlate with many others, and association can enhance or inhibit the activation state of adjoining neural units. Each individual neural unit figure out using summation function. There may be a threshold function or specify function on each connection and on the unit itself, such that the signal must pass the limit before proliferate to other neurons [5]. These systems are self-learning and trained, rather than obviously programmed, and excel in areas where the solution or feature detection is ambiguous to obvious in a classical computer business [6]. Fuzzy Logic (FL) is a method of reasoning that reassembles human reasoning. The approach of FL begin the way of decision catching in human that involves all intermediate possibilities between digital values yes and no [7]. Different techniques for data classification using Neuro-Fuzzy has been continually evolving to ensure efficient classification accuracy. In the literature survey there is different fuzzy method for classification. Fuzzy Neural Network Model [8] is divides into premise and consequence parts and then FUZNET structure is used. Improved Fuzzy Clustering Algorithm [9] divides into 3 features Improved Clustering Algorithm, Structure Identification, Parameter inference engine. Fuzzy Clustering Means Data Association Algorithm[10] gives high accuracy when target are far from each other , IT2FCM Based Neuro-Fuzzy Model[11] used karnik mendel model for type reduction, Adaptive Neuro-Fuzzy System Algorithm[12] are train by using hybrid network, Transductive Neuro-Fuzzy System Algorithm [13] used in sperical rolling robot for calculating fuzzy models, A Novel Neuro-Fuzzy classification Algorithm [14] used only 3 membership for each feature, Class-Based Grouping Algorithm used but by using this methods we cannot find classification with high accuracy. The output of Class-Based Grouping Algorithm in the form of cluster and mean and standard deviation. The output of Class-Based Grouping Algorithm is passed to Transition process to calculate fuzzy values. Output of Transition is forward to the Classification process. The Classification is done by using weka tool and used Multi-Layer Perceptron classifier to classify the dataset. After that EM Clustering Algorithm is used to increase the classification accuracy.

II. PROPOSED SYSTEM

In the below Fig. I the Abalone and Diabetes Dataset used as a input. Using this input performed the preprocessing, clustering and classification step. To preprocess the dataset we sorted the dataset in ascending order. After that used the Class-Based Algorithm for the purpose of Clustering and then Transition

Process and Classification Process is done. To reduce cluster size EM Clustering Algorithm is used.

2.1 Algorithm Used: Class-Based Grouping

The Dynamic Clustering Algorithm is also called as a “Class-Based Grouping” Algorithm. The Dynamic Clustering Algorithm works as follows [15]. Sort the dataset element in ascending order based on the similar class. Then group the element of the previous element and next element if the class is same then group that element and say it cluster. After that calculate centroid value. After that group the cluster based on threshold value. Then find out the centroid value and standard deviation value [15].

2.2 EM Clustering Algorithm

The target of this system is to assume some basic mean, standard deviation and pi values and iteratively enhance the estimate. Every single iteration is made up of two steps - the so *E* step and the *M* step [16]. The E step: This is the Expectation Part. Using the current mean and standard deviation assumption, we measure probabilities. We calculate these values for each feature value. This helps us forecast which class is responsible for which data point (called the responsibility).

$$r_{ic} = \frac{\pi_c N(x_i;\mu_c,\sigma_c)}{\sigma_{c|}\,\pi_{c|} N(x_i;\mu_{c|},\sigma_{c|})} \qquad (1)$$

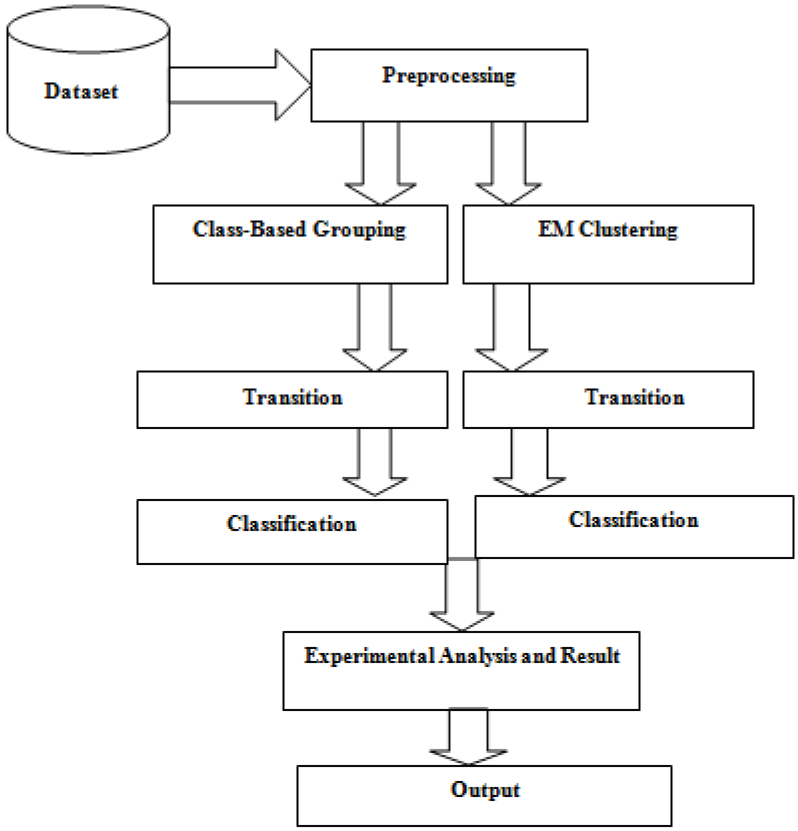


Figure.1.The Process Of Em Clustering For Classification Using Class Based Grouping Algorithm

Let's look at the numerator and denominator of this equation independently. The responsibility is calculated for each data point and for each one cluster. So is the responsibility of data point with regard to the class. This is represented by the numerator - which is just the probability of under the class. We multiply it by the weight as well - since this is a linear combination and we use to account for the constant factor [16].
Now let's look at the denominator. It is simply the sum of probabilities of the data point under all the class. This process, it acts as a balancing factor - ensuring that the value of is between zero and one for all classes. This establish that the value we get is a valid probability[17]. For calculating probability used the following formula:

$$p(x) = \frac{1}{\sigma * \sqrt{2\pi}}\, e^{-\frac{1}{2}}\langle\frac{x-\mu}{\sigma}\rangle^2 \qquad (2)$$

And that's it - the E step just associate computing this quantity called responsibility.

The M step: This is the Maximization Part. Using the responsibilities in the previous step, we renew the mean and standard deviations. Now that we have the responsibilities for every data point, we can use this to enhance our guess of each class mean, standard deviation and the weighting factor. To calculate the new mean for every cluster , we have:

$$\mu_c = \frac{1}{m_c}\sum_i r_{ic}\, x^{(i)} \qquad (3)$$

This is very similar to learning the mean of a single class. The idea is, you sum up all the points and divide them by the number of points. The only difference here is that we're using the "responsibility" to account for membership in a class.

In a single class, every data point is a member of the class - thus the responsibility is 1. However for multiple class, a single data point might contribute more to class A rather than class B. Thus responsibility accounts for this.

The denominator m_c takes into account this as well.

$$m_c = \sum_i r_{ic} \quad (4)$$

We have a very similar equation for calculating the standard deviation as well:

$$\sigma_c = \frac{1}{m_c} \sum_i r_{ic} \, (x^{(i)} - \mu_c)2 \quad (5)$$

Similarly, we can calculate the pi values of as well:

$$\pi_c = \frac{m_c}{m} \quad (6)$$

Here, m is the total number of data points in the dataset

Keep doing these two steps one after the other. Ultimately, the values of the mean and standard deviation will stabilize - that will be the learned model.

Transition Process

The number of clusters and initial values of the mean and standard deviation from a EM Clustering Process are used in the transition process. The original inputs are fed to the Gaussian membership function layer of the model. The number of membership functions for each feature is varies depending on the cluster of the feature.

$$\mu_{ij} \begin{cases} 0, if \ \sigma_{ij} = 0 \ and \ x_i \neq c_{ij}, \\ e^{\frac{1(x_i - c_{ij})2}{2\,\sigma_{ij}}}, if \ \sigma_{ij} \neq 0 \\ 1, if \ \sigma_{ij} = 0 \ and \ x_i = c_{ij}, \end{cases} \quad (7)$$

The membership values of feature *xi* . where *i* is identified the original feature and *j* is identified the cluster order of each feature. Normally the membership value calculated from Gaussian membership function. However the membership value is set to 0 if no distribution of that cluster and *xi* is not equal to mean of the cluster. The membership value is set to 1 if no distribution of that cluster and *xi* is equal to mean of the cluster. otherwise middle formula is used for calculating membership value.

Classification Process

The membership values are fed to the neural network. For classification Multi-layer Perceptron is used. The sigmoid function is used as activation in every node of the output layer. The output value can calculate from the following equation:

$$s_k = \sum_{i=1}^{m} \sum_{j=1}^{N_i} w_{kij} \, \mu_{ij} \quad (8)$$

$$o_k = \frac{1}{(1+e^{-s_k})} \quad (9)$$

In equation (8) k is the node in output layer, j is identified node in input layer and i is identified original feature. The w_{kij} is the weight of feature i that link between node k of output layer and node j of input layer. following is the example of EM Clustering Algorithm. The target of this system is to assume some initial mean, standard deviation and pi values and iteratively enhance the estimate. Every single iteration is made up of two steps - the so *E* step and the *M* step. The following is the 5 sample values of feature length of Abalone dataset.

Table.1. Em clustering example

Xi (Length)
4.40
4.75
4.83
5.01
4.98

Initialization: Initialize some basic mean, standard deviation and pi values

$\mu_1 = 1, \sigma_1 = 1, \pi_1 = 0.5$

Where μ_1 is mean σ_1 *is* standard deviation and π_1 is weighting factor.

E step: This is the Expectation Step. Using the current mean and standard deviation guess, we calculate probabilities. We calculate these values for each feature value (called responsibility). The probability of each feature value is calculated as follows by using following formula:

$$r_{ic} = \frac{\pi_c N(x_i; \mu_c, \sigma_c)}{\sum_{j\epsilon(0,k)} \pi_j N(x_i; \mu_j, \sigma_j)}$$

$N(x_i; \mu_c, \sigma_c)$ is calculated using following formula.

$$p(x) = \frac{1}{\sigma * \sqrt{2\pi}} \, e^{-\frac{1}{2}} \langle \frac{x-\mu}{\sigma} \rangle^2$$

$$p(4.40) = \frac{1}{1*\sqrt{2*3.14}} \, e^{-\frac{1}{2}} \langle \frac{4.40-1}{1} \rangle^2$$

$$= 3.425*10^{-3}$$

$$p(4.75) = \frac{1}{1*\sqrt{2*3.14}} \, e^{-\frac{1}{2}} \langle \frac{4.75-1}{1} \rangle^2$$

$$= 9.8*10^{-4}$$

$$p(4.83) = \frac{1}{1*\sqrt{2*3.14}} \, e^{-\frac{1}{2}} \langle \frac{4.83-1}{1} \rangle^2$$

$$= 7.25*10^{-4}$$

$$p(5.01) = \frac{1}{1*\sqrt{2*3.14}} \, e^{-\frac{1}{2}} \langle \frac{5.01-1}{1} \rangle^2$$

$$= 3.575*10^{-4}$$

$$p(4.98) = \frac{1}{1*\sqrt{2*3.14}} \, e^{-\frac{1}{2}} \langle \frac{4.98-1}{1} \rangle^2$$

$$= 4.03*10^{-4}$$

$$r_{ic} = \frac{\pi_c N(x_i; \mu_c, \sigma_c)}{\sum_{j\epsilon(0,k)} \pi_j N(x_i; \mu_j, \sigma_j)}$$

Denominator calculated as follows:

$$\sum_{j\epsilon(0,k)} \pi_j N(x_i; \mu_j, \sigma_j) = 5.8905 * 10^{-3}$$

Responsibility values for each feature as follows:

$$r_{4.40} = \frac{3.425 * 10^{-3}}{5.8905 * 10^{-3}} = 0.58 = 1$$

$$r_{4.75} = \frac{9.8 * 10^{-4}}{5.8905 * 10^{-3}} = 0.16 = 2$$

$$r_{4.83} = \frac{7.25 * 10^{-4}}{5.8905 * 10^{-3}} = 0.123 = 2$$

$$r_{5.01} = \frac{3.575 * 10^{-4}}{5.8905 * 10^{-3}} = 0.06 = 2$$

M step: This is the Maximization Step. Using the responsibilities in the previous step, we update the mean and standard deviations.

$$\mu_1 = \frac{1}{0.58}\sum_i r_{ic}\, x^{(i)}$$

$$= \frac{1}{0.58} * 0.58 * 4.40$$

$\boldsymbol{\mu_1}$**= 4.40** (mean of cluster 1)

$\boldsymbol{m_1 = 0.58}$ (Summation of responsibility that belong to cluster1)

$$\sigma_1 = \frac{1}{m_1}\sum_i r_{ic}\,(x^{(i)} - \mu_1)^2$$

$$= \frac{1}{0.58}(0.58(4.40 - 4.40)^2)$$

$\boldsymbol{\sigma_1}$ **=** ***0*** (Standard deviation of cluster 1)

$$\pi_1 = \frac{m_1}{m}$$

$$= \frac{0.58}{5}$$

$\boldsymbol{\pi_1}$ **=0.116** (Weighting factor of cluster 1)

$$\mu_2 = \frac{1}{0.403}\sum_i r_{ic}\, x^{(i)}$$

$$= \frac{1}{0.403} * 1.95349$$

$\boldsymbol{\mu_2}$**= 4.847 (**mean of cluster 2 as follows)

$m_2 = 0.16 + 0.123 + 0.06 + 0.06$

$\boldsymbol{m_2}$ **=0.403** (Summation of responsibility that belong to cluster 2)

$$\sigma_2 = \frac{1}{m_2}\sum_i r_{ic}\,(x^{(i)} - \mu_2)^2$$

$$= \frac{1}{0.403}[0.16(4.75 - 4.847)^2 + 0.123 * (4.834.847)^2 + 0.06 * (5.01 - 4.847)^2 + 0.06 * (4.98 - 4.847)^2]$$

$$= \frac{4.33*10^{-3}}{0.403}$$

$\boldsymbol{\sigma_2}$**= 0.010** (Standard deviation of cluster 2)

$$\pi_2 = \frac{0.403}{5}$$

$\boldsymbol{\pi_2}$ **=0.08 (**Weighting factor of cluster 2)

After that calculate the responsibility value of new cluster 1 and 2 and then again new clusters made then again recalculate the mean and standard deviation of new clusters. This step is is iterative until no change then we converge.

III. EXPERIMENTAL RESULT

The system is build using Java framework (version jdk 6) on Windows platform. Net beans (version 6.9) are used as a development tool. The system does not require any specific hardware to run; any standard machine is capable of running the application. The proposed system implemented on Abalone and Diabetes dataset. The information of Abalone and Diabetes dataset are as follows: The Abalone Dataset have the 8 different attributes and 3 different types of classes. The total size of dataset is 1510. The attributes and classes information of Abalone Dataset are as follows:

The Attributes are **Length:** Longest shell measurement, **Diameter:** perpendicular to length, **Height**: with meat in shell, **Whole weight:** whole abalone, **Shucked weight :** weight of meat, **Viscera_weight:** gut weight (after bleeding), **Shell_weight:** after being dried, **Rings**:+1.5 gives the age in years. The Classes are Male (M), Female (F), and Infant (I). The Diabetes Dataset has the 8 different attributes and 2 different types of classes. The total size of dataset is 768. The attributes and classes information of Diabetes Dataset are as follows: The Attributes are Number of times pregnant, Plasma glucose concentration a 2 hours in an oral glucose tolerance test, Diastolic blood pressure (mm Hg), Triceps skin fold thickness (mm), 2-Hour serum insulin (mu U/ml), Body mass index (weight in kg/(height in m)2), Diabetes pedigree function 8. Age (years). The Classes are Positive (1), Negative (0)

On the basis of following parameters Neuro-Fuzzy system are to be compared:

1. Accuracy
2. Mean Absolute Error

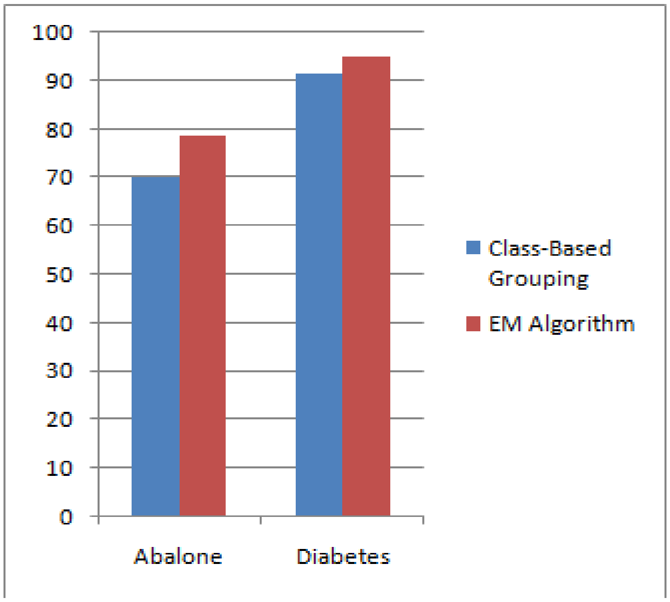
The Class-Based Grouping Algorithm and EM Algorithm is applied to Abalone and Diabetes dataset. The experimental result of both algorithm are as shown below. Table II shows the Accuracy and Mean Absolute Error of Class-Based Grouping and EM Algorithm for Abalone dataset. Table III shows the Accuracy and Mean Absolute Error of Class-Based Grouping and EM Algorithm for Diabetes dataset.

Table.2. Comparison of parameter using class-based grouping and em algorithm on abalone dataset

Evaluation Measures	**Class-Based Grouping Algorithm**	**EM Algorithm**
Accuracy	70.0	78.61
Mean Absolute Error	0.43	0.32

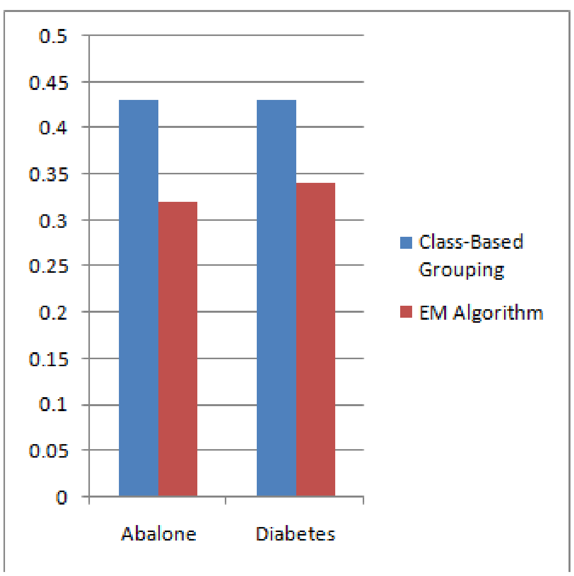
Table.3. Comparison of parameter using class-based grouping and em algorithm on diabetes dataset

Evaluation Measures	Class-Based Grouping Algorithm	EM Algorithm
Accuracy	91.32	95.10
Mean Absolute Error	0.43	0.34



Graph.1. Accuracy of class-based grouping and em algorithm

Graph I shows the Accuracy graph of both algorithm using Abalone and Diabetes Dataset. The Accuracy of EM algorithm is more on both dataset as compare to Class-Based Grouping Algorithm.



Graph.2.Mean absolute error class-based grouping and em algorithm

Graph II shows the Mean Absolute Error graph of both algorithm using Abalone and Diabetes Dataset. The Mean Absolute Error of EM Algorithm is less on both dataset as compare to Class-Based Grouping Algorithm.

IV. CONCLUSION

In Class-Based Grouping Algorithm the feature wise clusters of different classes are formed and based on it the membership function value of feature to the cluster is estimated. The cluster size is more by using Class-Based Grouping Algorithm. To reduce cluster size and improve the accuracy EM clustering is implemented. From the result of EM clustering reduces the clusters size as compare to Class- Based Grouping Algorithm. These algorithms are implemented on Abalone and Diabetes datasets and respective cluster size are recorded. A comparison of algorithm is done on the values of Accuracy, Mean Absolute Error and came to conclusion that EM Algorithm gives more accuracy as compare to Class-Based Grouping. By using Abalone Dataset the Accuracy of Class-Based Grouping is 70.00% while EM algorithm is 78.61%. Mean Absolute Error of Class-Based Grouping is 0.43 while EM algorithm is 0.32. By using Diabetes Dataset the Accuracy of Class-Based Grouping is 91.32% while EM algorithm is 95.10%. Mean Absolute Error of Class-Based Groupingis 0.43 while EM algorithm is 0.34.

V. REFERENCES

[1]. E. lughoter, “envloing fuzzy system: methodologies, advanced concept and application”, springer heidelberg , 2011

[2]. J. and J. Pei, “Data Mining: Concepts and Techniques”, Morgan Kauf-mann, San Francisco,. pp. 325-370, 2006.

[3]. McCulloch, Warren; Walter Pitts, “A Logical Calculus of Ideas Immanent in Nervous Activity”, Bulletin of Mathematical Biophysics. 5 (4): 115–133, 1943.

[4]. S. Mittal, “A Survey of Techniques for Approximate Computing”, ACM Computing Surveys, 2016

[5]. Hoskins, J.C.; Himmelblau, D.M., “Process control via artificial neural networks and reinforcement learning”, Computers & Chemical Engineering. 16 (4): 241–251, 1992

[6]. Cejas jesus, “compensatory fuzzy logic”, 2011

[7]. Petr Cermak and Pave1 Chmiel, “Parameters optimization of fuzzy- neural dynamic model”, 2004

[8]. Asli Celikyilmaz, “Enhanced Fuzzy System Models With Improved Fuzzy Clustering Algorithm”, IEEE transaction on Fuzzy Systems , vol.16,no. 3, June 2008.

[9]. Abdolreza Dehghani Tafti and Nasser Sadati, “Fuzzy Clustering Means Data Association Algorithm using an Adaptive Neuro-Fuzzy Network”, 2009

[10]. Emad Fallahzadeh and Mohammad Ali Montazeri, “Forecasting Foreign Exchange Rates Using an IT2 FCM based IT2 Neuro-Fuzzy System” , 2010

[11]. Erkan Kayacan, Herman Ramon, and Wouter Saeys, “Adaptive Neuro-Fuzzy Control of a Spherical Rolling Robot Using Sliding-Mode-Control-Theory-Based Online Learning Algorithm”, IEEE transaction on cybernetics, vol. 43, no. 1, February 2011

[12]. A. Gajate, R. E. Haber, P. I. Vega and J. R., “Alique A transductive neuro-fuzzy controller: application to a drilling process, Neural Networks”, IEEE Transactions, vol. 21, no. 7, pp. 1158-1167, July 2012

[13]. N. Eiamkanitchat, N. Theera-Umpon and S. Auephanwiriyakul, “A Novel Neuro fuzzy Method for Linguistic Feature Selection and Rule- Based

[14]. Classification,” The 2nd International Conference on Computer and Automation Engineering (ICCAE),pp. 247-252, 2013

[15]. Narissara Eiamkanitchat and Poonarin Wongchomphu, “Enhance Neuro-Fuzzy System for Classification using Dynamic Clustering”, The 4th Joint International Conference on Information and Communication Technology Electronic and Electrical Engineering (JICTEE), 2014

[16]. Amita Verma, Ashwani Kuma, “Performance Enhancement of K-means Clustering Algorithms for High Dimensional Data sets”, International Journal of Advance Research in Computer Science and Software Engineering, Volume 4 , Issue 1, January 2014

[17]. Garima Sehgal, Dr. Kanwal Garg,“ Comparison of Various Clustering Algorithms”, International Journal of Computer Science and Information Technology, Volume 5 , Issue 3, April 2014.

[18]. Weka: Data Mining Software in Java, http:// www. cs.waikato.ac.nz/m