A New Approach to Adaptive Neuro-Fuzzy Modeling using Kernel based Clustering

Sharifa Rajab[1] and Vinod Sharma[2]

[1] University of Jammu Campus

Received: 11 December 2014 Accepted: 5 January 2015 Published: 15 January 2015

Abstract

Data clustering is a well known technique for fuzzy model identification or fuzzy modelling for apprehending the system behavior in the form of fuzzy if-then rules based on experimental data. Fuzzy c- Means (FCM) clustering and subtractive clustering (SC) are efficient techniques for fuzzy rule extraction in fuzzy modeling of Adaptive Neuro-fuzzy Inference System (ANFIS). In this paper we have employed a novel technique to build the rule base of ANFIS based on the kernel based variants of these two clustering techniques which have shown better clustering accuracy. In kernel based clustering approach, the kernel functions are used to calculate the distance measure between the data points during clustering which enables to map the data to a higher dimensional space. This generalization makes data set more distinctly separable which results in more accurate cluster centers and therefore a more precise rule base for the ANFIS can be constructed which increases the prediction performance of the system. The performance analysis of ANFIS models built using kernel based FCM and kernel based SC has been done on three business prediction problems viz. sales forecasting, stock price prediction and qualitative bankruptcy prediction. A performance comparison with the ANFIS models based on conventional SC and FCM clustering for each of these forecasting problems has been provided and discussed.

Index terms— fuzzy modelling, kernel function, neuro-fuzzy model, fuzzy inference system, business prediction.

1 Introduction

he concept of fuzzy logic was introduced by Lofti Zadeh ??1969) based on fuzzy set theory in early 60's as an innovative approach to characterize the non-probabilistic uncertainty. since then this field has evolved into a productive realm encompassing various domains viz. fuzzy reasoning, fuzzy topology, fuzzy modelling and fuzzy inference systems. A Fuzzy inference system (FIS) is referred by a number of names like fuzzy model, fuzzy expert system, fuzzy associative memory and so on.

FIS is composed of three conceptual parts: a fuzzy rule base containing fuzzy rules, database defining the membership functions used in the fuzzy rules and a reasoning procedure for performing inference upon the rules and provided facts to obtain the output. The fuzzy if-then rules are used to represent the input-output relationships of the modeled system and are helpful to present the qualitative aspect of human reasoning without using any accurate mathematical model for the system. The fuzzy rule base for FIS can be constructed directly by domain experts a method that is usually error prone or by using fuzzy modelling approach. Fuzzy modelling also called fuzzy identification is an important technique to capture the behavior of a system to be modeled in the form of fuzzy if-then rules using its quantifiable characteristics. It has been addressed in a number of studies (Mamdani, 1976;Tong et al., 1980;Larsen, 1980) but was first discussed systematically by Takagi et al. (1985) as an effective technique for the estimation of dynamic fuzzy systems for problems of non linear uncertain nature.

The important issue of determining the number of fuzzy rules in the rule base and the values of parameters of membership function in fuzzy rules are dealt with using fuzzy modelling. Fuzzy modelling is nowadays successfully applied in control, prediction and other applications for the identification of fuzzy models using observed input output datasets. The fuzzy models possess the capability to provide insights into the relationships between various variables in the model which is not possible with several black box techniques such as neural networks. The fuzzy models also allow integrating the information obtained from the observed numerical input output data with the prior expert knowledge. A standalone FIS however does not have the ability to learn and can be extended by using optimization and adaptive methods for performance improvements. ANFIS proposed by ??ang (1995) is a widely employed fuzzy model based on the concept of integrating fuzzy inference systems and neural networks that uses learning to fine tune its fuzzy rule base for optimizing the system inference process. It combines the human like reasoning method of fuzzy systems based on fuzzy rules with the learning capability and connectionist structure of neural networks.

For fuzzy modelling different data partitioning techniques like clustering such as FCM clustering are used to obtain the partitions of the dataset to capture the internal trends of the input output data samples. These partitions or clusters are then used to construct the fuzzy if-then rules for the fuzzy model being build and then a method is used to fine tune the initial rule base to obtain the final rule base. The fuzzy modelling of ANFIS based on measured input-output data is generally performed using one of the three methods namely Grid Partitioning, Subtractive clustering and FCM clustering. The grid partitioning although an efficient method to partition the input space, has some disadvantages like curse of dimensionality of input, computation cost, exponential expansion of rule base etc. Data clustering is a handy alternative technique for fuzzy modelling where the clusters obtained from a clustering algorithm are used as a basis for fuzzy rule generation. Fuzzy identification using clustering consists of finding the clusters in the data space and using the obtained cluster centers to calculate the premise and consequent parts of the fuzzy rules. Therefore the accuracy of clustering determines the quality of the rule base and hence the performance of the resulting fuzzy model. Both the approaches viz. SC proposed by Chiu (1994) and FCM clustering technique proposed by Dunn (1974) and later improved by ??ezdek et al. (1984) are efficient techniques for clustering data sets and hence effective modelling techniques. Recently, the kernel methods have achieved popularity for various classification and regression based problems. The accuracy of SC and FCM clustering is improved by incorporating kernel functions in the calculation of the distance measures between the data points during clustering process which results in more precise cluster centers (Kim et al., 2004; ??iang et al., 2004). Higher clustering accuracy is achieved as the kernel induced distance measures increase the data separability by using the higher dimensional space which reveals more precise data partitions.

In this paper we have a proposed novel modelling techniques for the fuzzy rule base construction of ANFIS based on kernel based SC (KSC) and kernel based FCM (KFCM) techniques. To build the prediction model the data set is first partitioned into clusters using kernel based clustering, the resulting cluster centers are then employed to build the initial fuzzy rule base for ANFIS and then the resulting rule base has been optimized using a hybrid learning algorithm consisting of standard Backpropagation and least square estimation. The effectiveness of the KSC and KFCM based ANFIS models has been tested on three business prediction problems namely qualitative bankruptcy prediction, sales forecasting and stock price prediction. A comparison with the ANFIS models based on original SC and FCM for these business prediction problems has also been presented. This paper organization is as following. Section 2 deals with the review of the prior relevant research. In section 3 the research methodology is presented that gives the details of FCM and KFCM algorithms, provides an overview of ANFIS and fuzzy rule generation methods based on clustering and also presents and discusses the simulation results providing a comparison of performance with the ANFIS based on conventional FCM clustering and SC. Section 4 provides the concluding remarks on this study and various enhancements to this work.

2 II.

3 Previous Work

Recently a number of studies (Yao et al., 2000; ??ejan et al., 2011;Hossein et al., 2010;Kalhor et al., 2009; ??uk et al., 2003) have addressed the problem of fuzzy model identification based on the data clustering algorithms. In several studies the kernel methods have also been employed along with various conventional clustering techniques for this purpose. Yang et al. ??2008) proposed a novel method for fuzzy modelling of a Takagi-Sugeno system based on dual kernel-based method. The authors used a conventional FCM algorithm for partitioning data into various clusters. Then a kernel function independent of the parameter selection problem was used to locate the support vectors within each of the clusters. The experimental results from the study showed that the method lead to a fuzzy model with concise structure having good generalization capability. Also the performance of the system was not affected by the initial cluster number needed in FCM. Lukasik et al. (??008) presented a kernel-density gradient estimation technique for fuzzy rule extraction. The authors used clustering based on kernel density estimator. The cluster centers obtained on clustering were then used to construct the rule base. The assumption underlying the technique was that local maximum of a kernel estimator of a probability density function for m-dimensional data can be used as a basis for each cluster. But instead of the density function the authors used the gradient of the density function in cluster center identification. The neuro-fuzzy system based on this approach was experimented for non-linear function approximation and controller synthesis and showed good performance. Suga et al. (**??**006) used an iterative feature vector selection (FVS) based on kernel method to calculate membership function parameter values and the number of fuzzy rules for a Takagi-Sugeno fuzzy model. The kernel based FVS algorithm was used to obtain a basis of data space called feature vector into the feature space. This feature vector was then used as the center of a membership function in the antecedent part of the fuzzy rule. After finding the premise parts of the fuzzy rules, the coefficients of the consequent of the fuzzy rules were obtained using least square methods. The proposed system was applied to the modelling of a two input non-linear function which showed the effectiveness of the proposed fuzzy system for non-linear system modelling. Almost all of the above research studies proved that the kernel methods can be used to enhance the performance of the fuzzy rule based models but the use of KSC and KFCM for fuzzy modelling of popular ANFIS model was not explored which is undertaken in this paper.

4 III.

5 Methodology

This study deals with the performance analysis of ANFIS built using kernel based clustering techniques viz. KSC and KFCM for business prediction problems. The prediction model for a problem is built in three stages: 1) dataset is partitioned into various clusters using one of these kernel based clustering techniques, 2) the cluster centers obtained from clustering are used to build the fuzzy rule base of ANFIS, and 3) the resulting ANFIS model is trained using the hybrid learning algorithm consisting of gradient descent method and least square method. The various techniques used have been discussed in the following sections.

6 a) Techniques employed i. Kernel based subtractive clustering

The original SC is based on calculating the potential function called mountain value at each data point. It is an improved version of the mountain method and uses each input data point in the dataset as a potential cluster center rather than using grid based formulation in mountain clustering method thus leading to lower computational complexity for higher dimensional data sets. KSC was proposed by Kwang et al. (**??**004) as an improvement over conventional SC algorithm where kernel functions are employed in potential value calculation. In original subtractive clustering, for a dataset X={x 1 , x 2 , ?,x n }, the potential value at each data point x i is given by:? e ??||x i ?x j || 2 n j=1 , ? = 4 r a 2 (1)

where r a is a positive constant called cluster radius defining the range of influence of a cluster center along each data dimension and affects the number of clusters generated. The data point with the highest potential value P 1 is selected as the cluster center ?? 1 . In order to find the subsequent cluster centers using the same procedure the potential value for each data point x i is modified as:P i = P i ? P 1 * e ??||x i ?c 1 || 2 , ?= 4 r b 2 , r b= ?r a (2)

Here r b is a positive constant and ? is the squash factor used to squash the potential values for the distant points to be considered as part of a cluster. It is evident that the reductions in potential values of data points near the newly found cluster center is more than the distant points and hence have a least chance of being selected as cluster centers.

Using kernel approach the kernel functions are employed in calculating the distance measure given by: ||x i ? x j || 2 and ||x i ? c 1 || 2 in Eqs. (**??**) and (**??**) so that the data points are mapped to a higher dimensional space which makes the dataset more distinctly separable resulting in more informative potential values. Therefore, the centers produced are more accurate and when used in fuzzy modelling can result in more useful fuzzy rule base for a fuzzy mode like ANFIS.

The basic notion in kernel methods is a nonlinear mapping Ø to a higher dimensional space from the input space i.e. for a dataset X={x 1 , x 2 , ?, x n }: Ø: x Ø(x) (3) Using this non-linear mapping the dot product 'x i .x j ' used as a similarity measure in various learning algorithms can be mapped to a more general measure: Ø(x i).Ø(x j). This dot product in higher dimensional space is calculated using a kernel function K(x i ,x j) i.e.:

Ø(x i).Ø(x j) = K(x i ,x j) (4) The distance measure ||x i ? x j || 2 in input space in terms of function Ø therefore is given by:||x i ? x j || 2 = ||Ø(x i) ? Ø(x j)|| 2**(5)**

where:||Ø(x i) ? Ø(x j)|| 2 = (Ø(x i) -Ø(x j)) -(Ø(x i) -Ø(x j)) = Ø(x i).Ø(x i) -2Ø(x i) Ø(x j) + Ø(x j).Ø(x j) = K(x i ,x i) -2 K(x i ,x j) + K(x j ,x j) (6)

Thus for KSC eq. (**??**) can be altered to incorporate kernel function by using eqs. (5) and (**??**):? e ???K?x i , x i ??2K?x i , x j ?+K?x j , x j , ?? n j=1**(7)**

K(x, y) can be any kernel function like gaussian kernel, polynomial kernel, fisher kernel etc. After a data point is selected as a cluster center in KSC, the potential function of other data points to find the subsequent centers is calculated as:P i = P i ? P * * e ?? ? K?x i , x i ??2K?x i , x * ?+K(x * , x *) ? (8)

where ?? is the positive constant in eq. (**??**) and x* is the newly obtained cluster center with potential value ?? * . After revising the potential of other data points, the data point with the highest potential is chosen as the second cluster center and the potential values of other data points are changed as in eq. (**??**). In general when nth cluster center ?? ?? * is selected, the potential of other data points is revised as: where ??? is the Accept Ratio i.e. a threshold potential value below which the data point is rejected as the cluster center and ?? is the

Reject ratio which specifies a threshold potential above which the data point is definitely accepted.Pi=P i -P n * * e ???K?x i , x i ??2K?x i , x n * ?+K(x n * , x n *)? (9)

ii. Kernel based FCM clustering KFCM was proposed by Qiang (2004) as an enhancement of the standard FCM clustering algorithm based on the use of kernel functions. For a dataset X={x 1 , x 2 ,?, x n }, the conventional FCM algorithm calculates the fuzzy subsets of X by minimizing an objective function given by:? ? ? ij m ||x j ? v i || 2 n j=1 c i=1 (10)

where n is the number of data points, c is the number of cluster centers, µ ij is the membership of x j in ith class, v i is the ith cluster center and m is the quantity to control the fuzziness of clustering. In KFCM the distance measure is generalized by employing a non linear mapping Ø from input space to a higher dimensional space i.e.:

Ø: x Ø(x) Therefore, as in KSC using (4) and (5) by the kernel approach the objective function in KFCM is given by:J m (U,V) = ? ? ? ij m n j=1 c i=1 ||Ø(x j) ? Ø(v i)|| 2**(11)**

From eq. (**??**):||Ø(x j) ? Ø(v i)|| 2 = K(x j , x j) -2 K(x j , v i) + K(v i , v i)**(12)**

K(x, y) can be any kernel function for example gaussian kernel, polynomial kernel, fisher kernel etc.

Using equation (12) eq. (11) becomes:? ? ? ij m n j=1 c i=1 (K(x j , x j) ? 2 K(x j , v i) + K(v i , v i)) (13)

Gaussian function is a common kernel function given by: K(x, y) = e (??|x?y|? 2 /? 2) (14) where K(x, x) = 1 and ? is an adjustable parameter. Using gaussian kernel function the eq. (**??**3) becomes:? ? ? ij m n j=1 c i=1 (1 -K(x j , v i))**(15)**

Whereµ ij = (1/(1?K(x j ,v i))) 1/(m ?1) ? (1/(1?K(x j ,v k))) 1/(m ?1) c k =1**(16)**v i = ? ? ij m K(x j ,v i) x j ?? ?? =1 ? ? ij m K(x j ,v i) ?? ?? =1**(17)**

Other kernel functions can also be used so that above equations can be modified accordingly.

7 Algorithm for KFCM

Step 1: set k=0, m > 1 and ?? > 0 for some positive constant.

Step 2: initialize the memberships ? ij 0 .

Step 3: I) Update all v i k using eq. (17).

II) Update all ? ij k using eq. (16).

8 If max ???

ij k ? ? ij k?1 ? ? <= ?? Stop else k=k + 1 go to step 3. end if

The fuzzy inference system has the capability of a non-linear system being modeled in terms of fuzzy if then rules. The fuzzy model identification therefore involves the determination of parameters for the premise membership functions and parameters in the consequences. Applying the clustering algorithm on the experimental dataset for the system to be modeled each of the resulting cluster centers essentially is an exemplary data point representing the system's characteristic behavior. Therefore, using clustering for fuzzy modelling each of the cluster centers is considered as a basis for a fuzzy rule for the initial rule base of the fuzzy inference system being modeled. Hence the number of cluster centers generated determines the number of the fuzzy rules for the modeled system.

The fuzzy model identification using data clustering techniques has been addressed in a number of studies **??**Han et Babuska et al. (1994) provides an effective method for fuzzy rule generation from the FCM generated fuzzy clusters where premise membership functions are obtained using projection of fuzzy clusters which can be orthogonal or eigenvector projection. The consequent parameters using this method can be obtained using least square estimation. According to Degado et al. (1997) the antecedent and consequent parameters can be directly obtained from cluster centers instead of projections on domains of outputs and inputs.

A number of studies on fuzzy modelling using subtractive clustering have used the method presented in paper (Chiu 1994). With this method if k cluster centers {c 1 ,?, c k } are generated in m-dimensional space, each of the vector c i can be decomposed into two vectors X i and Y i where X i represents the first n elements of c i corresponding to the input variables and Y i contains m-n output variables. For an input vector x the degree of fulfillment of rule i is given by:? i (x) = e ??||x?X i || 2

(18) Where ? is the positive constant used in eq. (**??**). The output vector y can be computed as:y = ? Y i ? i k i?1 ? ? k i=1 i**(19)**

A typical fuzzy rule has the following form: If x 1 is A 1 and ? and x n is A n then y 1 is B 1 and ? and y n is B n Where x i is the ith input variable, y i is the ith output variable and A i is the ith antecedent membership function and B i is singleton. Each of the ith rule is determined by the cluster center c i and each rule has multiple input variables and hence membership functions. If A j is the jth membership function of a rule i, it is given by:A j (x) = e ??(x?X ij) 2

And the consequent B i is given by: B j = Y ij (21) Where X ij is the jth element of X vector and Y ij is the jth element of Y vector of center c i . This method of rule generation achieves significant accuracy if the Takagi-sugeno type fuzzy rules are used in which the consequent parameters are the linear combination of input variables (Chiu 1994).

iv. ANFIS architecture ANFIS is an adaptive system that has the learning capability to optimize the performance based on finding the best parameters for the fuzzy rules within its rule base. Fig. 1 shows the architecture of ANFIS with two inputs x 1 and x 2 and a rule base consisting of consisting of two Sugeno type fuzzy rules:

If x 1 is A 1 and x 2 is B 1 then f 1 = p 1 x 1 + q 1 x 2 + r 1 If x 1 is A 2 and x 2 is B 2 then f 2 = p 2 x 1 + q 2 x 2 + r 2

The details of the functioning of each layer of the ANFIS are as follows: Layer 1: This is the input layer and consists of nodes with adaptive node functions. Each node has an output equal to:O 1,i = A i (x) f or i = 1,2**(22)**

Here output of each node is the value of the membership function A of that node and O k , I is the node in the i-th position of the k-th layer.

Various types of membership function can be used, like gauss function, the bell-shaped function etc. Layer 2: In this layer each node computes the product of incoming signals with output given by:O 2,i = w i = ?A i ?B i (y) , i = 1,2**(23)**

Layer 3: In this layer each j-th node computes the ratio of the firing strength of the j-th rule and the sum of all the firing strengths, with output:O 3,j = w j = w j w1+w2 , j = 1,2**(24)**

Layer 4 : In this layer function for i-th node is:O 4,1 = w ? i f i = w ? i (p i x + q i x + r i)**(25**

) Layer 5 : This layer has a single node that computes the overall output as the sum of all incoming signals:O 5,1 = ? w ? i i f i = ? w i f i i ? w i i (**26**)

Where O 5,1 is the obtained output available to user.

For the optimization of the fuzzy rule base of ANFIS either standard back propagation or the hybrid learning algorithm can be used. The hybrid learning is mostly used and is an effective technique which uses gradient descent method to update the premise parameters of the fuzzy rules and LSE is used to identify the optimal consequent parameters. In order to test the effectiveness of the KSC and KFCM based ANFIS models we used three datasets one for each of the three business prediction problems viz. qualitative bankruptcy prediction, sales forecasting and stock price prediction. These problems were selected as these are the popular research problems in business field nowadays and ANFIS model has been extensively applied to these problems in numerous studies successfully. For the purpose of qualitative bankruptcy prediction the dataset has been collected from one of the largest banks in Korea consisting of 260 services and manufacturing companies for period 2001-2002. This dataset has also been used by Jong et al. **??**2003) to discover the bankruptcy decision rules based on experts, decisions using genetic algorithm. Half of the companies in this dataset are bankrupt and other half non-bankrupt according to the classification done by the experts having an experience of nine years in this area. This dataset is based on six qualitative risk factors as listed in fig. 1 (a). Each of the factors is assigned an appropriate level viz. positive (P) or negative (N) or average (A). The output is the class of the company i.e. bankrupt (B) and non-bankrupt (NB) as shown in Fig. 1 (b).

The dataset used in the stock price prediction is the daily BSE stock data obtained from Yahoo finance for a period eight years from 1/2/2007 to 30/12/2014 consisting of 1966 records. The dataset is composed of five fundamental stock quantities (open price, maximum price, minimum price, stock trading volume and close price). We have used 70% consisting of 1179 records of this dataset for training, 20% consisting of 394 records as checking data and rest 20% consisting of 394 records as testing data for all the ANFIS models.

For sales forecasting problem we have used the sales data of chocolate items of a major distributor in Jammu city (India) collected for a period of five months from 1/12/2014 to 30/12/2015 consisting of 150 records. The dataset has four attributes viz. present day sale amount, maximum daily temperature, minimum temperature and next day sale. The temperature attributes have been included as the sale of chocolate items is affected by the temperature during a period.

9 c) Experimental results

In this section simulation results of the application of the KSC and KFCM based ANFIS models for qualitative bankruptcy prediction, sales forecasting and stock price prediction have been presented. In all the experiments the performance of the KFCM and KSC based ANFIS models has been compared with the conventional FCM and SC based ANFIS models respectively. For all the ANFIS models online learning has been used with an initial step size 0.01 and gaussian membership function given by:?? ? 1 2 ? ????? ?? ? 2 (27)

has been used for input variables in the first layer of ANFIS . The parameter c in eq. (**??**7) is the center of membership function, ?? determines the width of the membership function and x is the input variable. The values used for various parameters explained in section 3.1.1 for both the SC and KSC algorithms were: accept ratio = .5, reject ratio = .15 and squash factor = 1.25 for all the simulation examples. For all the experiments we have used gaussian kernel function defined in eq. (**??**3) as the kernel for the implementation of both KSC and KFCM technique. All the experiments were performed in MATLAB R2013a environment.

i. Qualitative bankruptcy prediction The first business prediction problem considered is the qualitative bankruptcy prediction. Out of the total 260 records of the dataset used for this problem, 75% has been used for training, 15% as checking data and 20% as testing data for both KSC and KFCM clustering based ANFIS models.

The cluster radius r a defined in eq. (**??**) for subtractive clustering takes values between 0 and 1 and strongly affects the number of clusters generated. A large value for r a results in lesser number of clusters and therefore lesser fuzzy rules in rule base and vice versa. The number of fuzzy rules in the system in turn affects the forecasting performance of the system. Therefore finding the optimum value for r a is important for a problem under consideration.

A value of .5 for parameter r a for both KSC and SC produced the ANFIS systems with lesser training, checking and testing errors and therefore was optimum for this dataset. KSC resulted in 24 clusters so that resulting ANFIS system had 24 fuzzy rules with 5 gauss membership functions associated with each input variable. The SC resulted in 49 clusters and the ANFIS system based on it had 49 rules. Therefore, KSC resulted in a less complex ANFIS with lesser number of parameters to be optimized than ANFIS based on SC. Fig. 2 shows the training and checking RMSE for both the KSC and SC based ANFIS models. After the systems were trained for 150 epochs for this dataset the root mean square error (RMSE) for both the systems did not change significantly. Fig. 2 The number of clusters used as the input parameter to the FCM and KFCM algorithms results in an equal number of fuzzy rules for the ANFIS system to be implemented which in turn considerably affects the performance of the system. Both the FCM and KFCM based ANFIS systems containing three fuzzy rules with three membership functions of gaussian type for each input attribute showed the lowest RMSE. On training for 50 epochs the RMSE for both the systems remained constant. Fig. 2 shows the training and checking RMSE curves for both the systems where it is evident that KFCM based ANFIS is better than FCM based one for this problem.

A value of 150 was used for the adjustable parameter ?? of gaussian kernel defined in (13) for both KSC and KFCM so that the resulting ANFIS models gave the lowest training, checking and testing errors.

ii. Sales Forecasting The next simulation example is the sales forecasting based on the sales data collected by authors. 75% of this dataset containing total 150 records has been used for the training the models, 15% for checking and 25% for testing.

With parameter r a = .5, the KSC resulted in 4 clusters for this dataset. The resulting ANFIS system contained four fuzzy rules with four membership functions of gaussian type associated with each input. The SC with r a = .5 resulted in 2 fuzzy rules for this dataset so that the resulting SC based ANFIS contained two fuzzy rules in rule base. In case of KSC values near 10 for the adjustable parameter ?? of gaussian kernel defined in (13) **??**.

10 iii. Stock price prediction

We have used the KFCM and KSC based ANFIS models for predicting the close price of the day based on the daily open price, trading volume, maximum and minimum stock price.

With r a = .5 the KSC based ANFIS system has two fuzzy rules with two membership functions associated with each input and the conventional SC based system has three fuzzy rules with three membership functions associated with each input variable. Both the systems were trained for 200 epochs after which the RMSE remained constant. The training and checking RMSE curves for both the ANFIS systems are provided in fig. **??**. After testing the KSC based ANFIS we had RMSE = .0034 and APE = 1.041%. For SC based ANFIS we had RMSE = .0037 and APE = 1.3961%.

The KFCM and FCM algorithm with 10, 5 and 3 cluster centers were used resulting in ANFIS systems with 10, 5 and 3 fuzzy rules respectively. But the FCM and KFCM based ANFIS containing 3 fuzzy rules were found to give the lowest errors. Fig. **??** shows the training and checking RMSE curves for these systems each containing 3 fuzzy rules. On testing the KFCM based ANFIS system resulted in RMSE = .0034 and APE = .7060% and the FCM based system showed RMSE = .0035 and .7471%.

For both KFCM and KSC algorithms a value of 150 for parameter ?? for gaussian kernel function resulted in ANFIS systems with lowest errors.

11 IV.

12 Conclusion

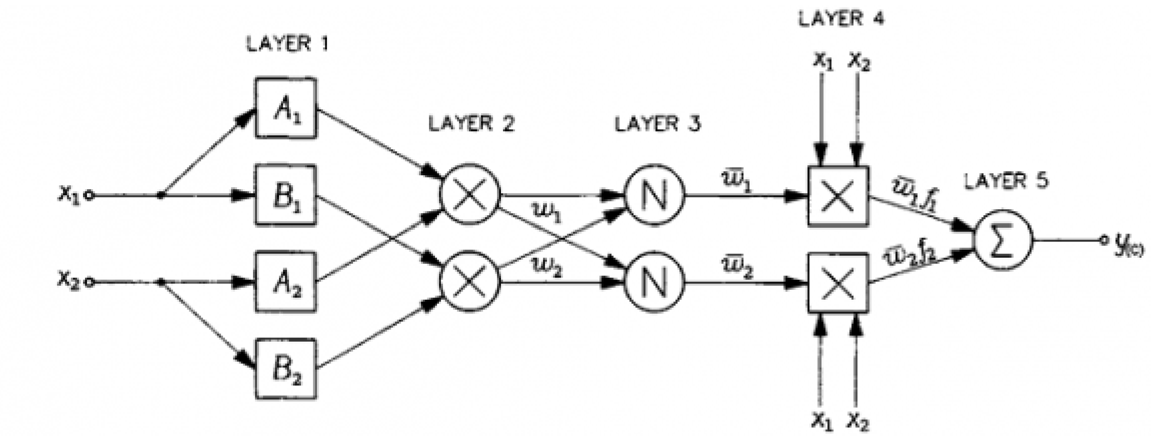
algorithms for extracting rules for initial fuzzy rule base for popular neuro-fuzzy model ANFIS. We used the kernel clustering based ANFIS models for three well known business prediction problems. For all the experiments the kernel based methods resulted in optimum number of fuzzy rules in the rule base of ANFIS, giving a lesser complex system. Moreover, the performance of these systems was mostly better in terms of training, checking and testing errors than the ANFIS models based on conventional subtractive and FCM clustering methods for these forecasting problems.

In this study we have used the gaussian kernel function for all the experiments. The major issue in using the kernel methods is to select the kernel function to be used, this work can be extended by using the multiple kernel based clustering approach which can overcome the problem of selecting the best kernel function for a particular data set. A performance comparison between the KFCM and KSC based ANFIS may be explored. Furthermore we have only considered ANFIS model which is currently most popular neuro-fuzzy system but other fuzzy models like type-2 fuzzy models can also be considered. Grid partitioning is also used for fuzzy modelling in ANFIS and gives satisfactory prediction accuracy but in this study we have not compared the performance of the kernel clustering based ANFIS models with such systems. The conventional subtractive and FCM clustering techniques have been successfully used for fuzzy modelling. But the kernel based variations of these algorithms have shown better clustering accuracy by using higher dimensional space which results in better data space partitioning.

Therefore when used in fuzzy modelling these techniques can result in a more useful fuzzy rule base for a fuzzy logic basis system. In this paper we have used the kernel based variants of these

13 References References Referencias

Figure 1:



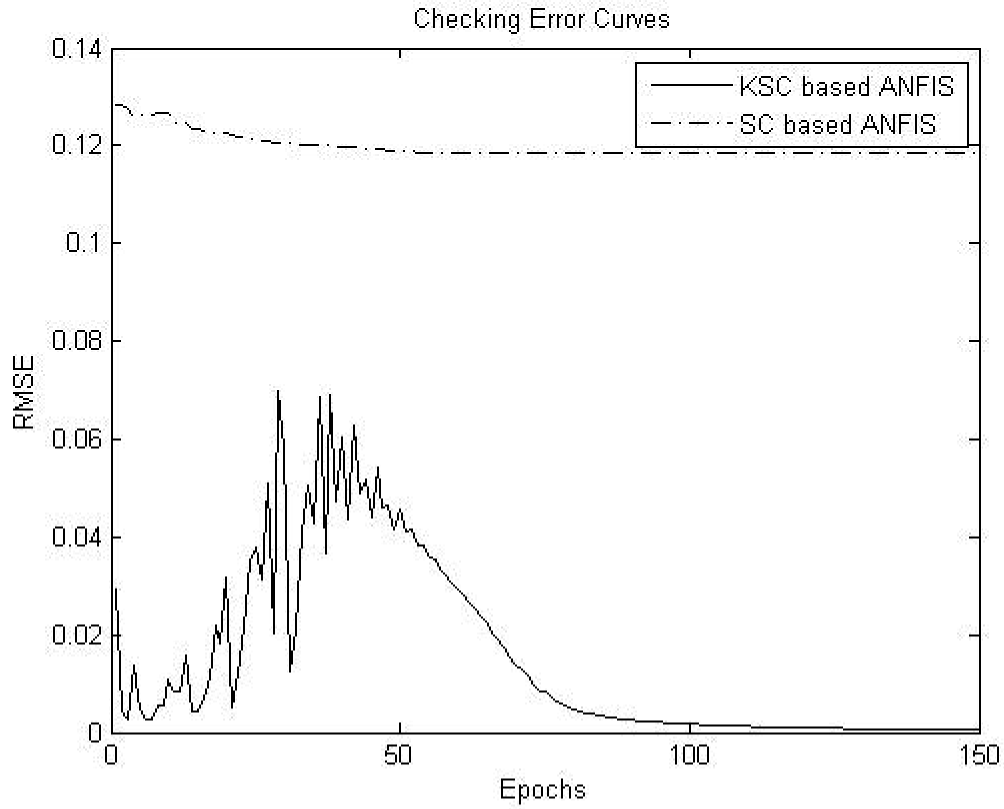
1

Figure 2: Fig. 1 :

[1] [2]

[1] © 2015 Global Journals Inc. (US) 1

[2] © 2015 Global Journals Inc. (US)



2

Figure 3: Fig. 2 :

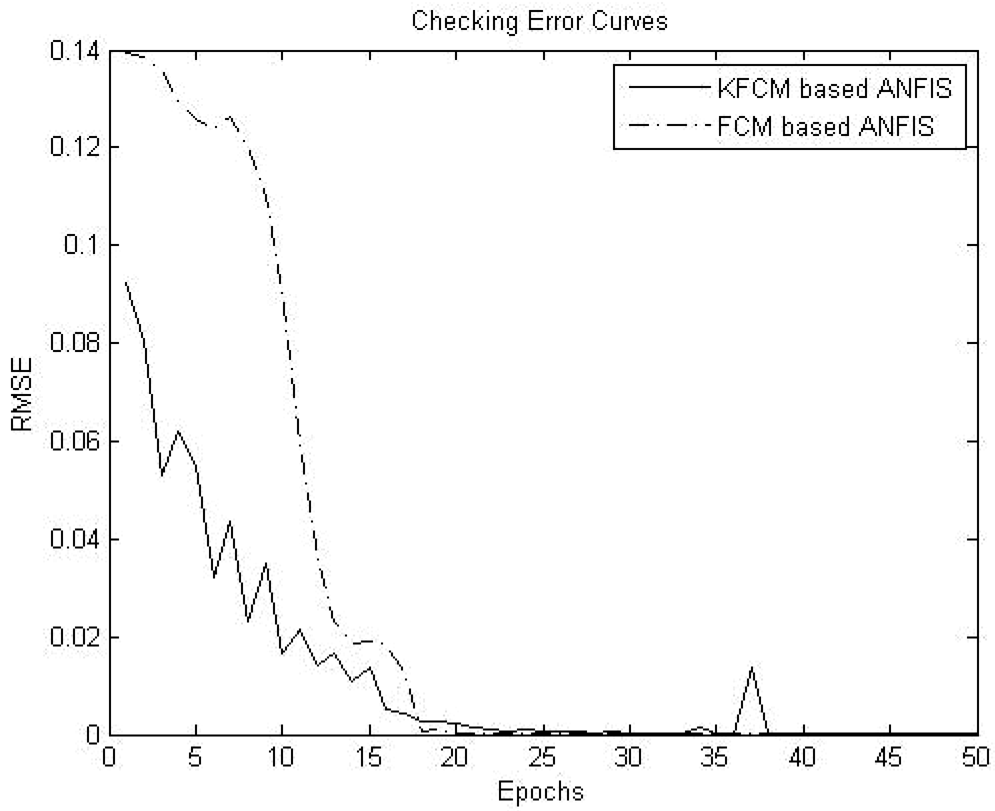


Figure 4:

accept x n

center and repeat

else if P n * < ? P 1 * reject x n

clustering process

else

let d m = least distance between x n

all earlier found cluster centers

if ?? + P n P 1 * * >= 1 accept x n

??

??

??

center and repeat clustering

else reject x n * and set potential

end if end if P n with the next highest potential * = 0. Select the data point as the cluster center

( )

Global Journal of C omp uter S cience and T echnology

[Note: D]

Figure 5:

1

30

Figure 6: Table 1 :

[James et al.] , C James , Robert Bezdek , Ehrlich . William Full MEANS CLUSTERING.

[Yao et al. ()] '17. fuzzy c-means clustering for recursive fuzzy identification of time-varying processes'. J Yao , M Dash , S T Tan , H Liu . *Fuzzy Sets and Systems* 2000. Elsevier. 113 p. . (ISA Transactions)

[Kim et al. ()] *A kernel-based subtractive clustering method Pattern Recognition Letters*, Dae-Won Kim , Ki Young Lee , Doheon Lee , H Kwang , Lee . doi:10.1016/j. patrec. 2004.10.001. 2004. Elsevier.

[Kim et al. ()] 'A new approach to fuzzy modelling of Nonlinear dynamic systems with noise, relevance Vector learning mechanism'. Jongcheol Kim , Y Suga , Sangchul Won . doi: 10. 1109 / TFUZZ.2005.864083. *IEEE Transactions on Fuzzy System* 2006.

[Li and Yang ()] 'A new approach to TS fuzzy modelling using dual kernelbased learning machines'. Wei Li , Yupu Yang . *Neurocomputing* 2008. 2008. 71 p. .

[Babuska and Verbruggen ()] 'A new identification method for linguistic fuzzy models'. R Babuska , H B Verbruggen . *at Open Meeting of the FALCON Working Group*, (Aachen) 1994.

[Jang ()] 'ANFIS : Adaptive-Network-Based Fuzzy Inference System'. Jyh-Shing Roger Jang . *FCM: THE FUZZY c-ALGORITHM*, 1993. 1984. 23 p. . (Computers & Geosciences)

[Mamdani ()] 'Application of fuzzy algorithms for control of simple dynamic plant'. E H Mamdani . *IEEE Proceedings of Institution Electrical Engineers* 1976. 121 p. .

[Celikyilmaz and Turksen ()] 'Enhanced Fuzzy System Models With Improved Fuzzy Clustering Algorithm'. A Celikyilmaz , I B Turksen . 10.1109/TFUZZ.2007.905919. *IEEE Transactions on FuzzySystems* 2007.

[Tong et al. ()] 'Fuzzy control of the activated sludge wastewater treatment process'. R M Tong , M B Beck , A Latten . *Automatica* 1980. 16 p. .

[Takagi and Sugeno ()] 'Fuzzy identification of systems and its applications to modelling and control'. T Takagi , M Sugeno . *IEEE Trans. Syst* 1985. 15 p. . (Man, Cybern.)

[Fuzzy model identification based on cluster estimation Journal of Intelligent and Fuzzy Systems] 'Fuzzy model identification based on cluster estimation'. *Journal of Intelligent and Fuzzy Systems* IOS press. 2 p. . (Stephen chiu 1994)

[Kim et al. ()] 'Fuzzy Model Synthesis with Kernel-Density-Based Clustering Algorithm'. Jongcheol Kim , Y Suga , Sangchul Won . 10.14864/softscis.2006.0.1934.022. doi: 10.1109/FSKD.2008.139. *Fifth International Conference on Fuzzy Systems and Knowledge Discovery*, 2006. 2008. Japan Society For Fuzzy Theory and Intelligent Informatics. (New Approaches to Fuzzy Inference System Using Kernel Machines)

[Han et al. ()] 'Fuzzy Models Synthesis with Kernel-Density-Based Clusterin Algorithm'. Pu Han , Jian-Zhong Shi , Dong-Feng Wang , Song-Ming Jiao . 10.1109/ICMLC.2010.5580478.15. DOI:10.1109/FSKD.2008.139. *FCM clustering algorithm for T-S Fuzzy model identification, International Conference on Machine Learning and Cybernetics*, 2010. 2008. (Fifth International Conference on Fuzzy Systems and Knowledge Discovery)

[Larsen ()] 'Industrial application of fuzzy logic control'. P M Larsen . *Int. J. Man-Machine Studies* 1980. 12 p. .

[Kalhor and Lucas ()] 'Online Identification of a neuro-fuzzy model through indirect fuzzy clustering of data space'. A Kalhor , C Lucas . 10.1109/FUZZY.2009.5277139. *IEEE International Conference on Fuzzy Systems*, 2009.

[Soleimani et al. ()] *Recursive Gath-Geva clustering as a basis for evolving neuro-fuzzy modeling, Evolving Systems*, Hossein Soleimani , B , Caro Lucas , N Babak , Araabi . 2010. Springer-Verlag. 1 p. .

[Degado et al. ()] 'Some Methods to Model Fuzzy Systems for Inference Purposes'. M Degado , A F Gomez , F Martin . *International Journal of Approximate Reasoning* 1997. 16 p. .

[Dunn ()] 'Some recent investigations of a new fuzzy partition algorithm and its application to pattern classification problems'. J C Dunn . *J. of Cybernetics* 1974. 4 p. .

[Myoung et al. ()] 'The Discovery of experts' decision rules from Qualitative bankruptcy data using genetic algorithms'. Jong Myoung , Ingoo Kim , Han . *Expert Systems with Applications* 2003. Elsevier. 25 p. .

[US) Guidelines Handbook Global Journals Inc ()] 'US) Guidelines Handbook'. `www.GlobalJournals.org` *Global Journals Inc* 2015.

[Zadeh ()] L A Zadeh . *Fuzzy sets, Information and Control*, 1965. 8 p. .