Neural Networks 104 (2018) 1–14

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

Design of double fuzzy clustering-driven context neural networks

Eun-Hu Kim [a], Sung-Kwun Oh [a,*], Witold Pedrycz [b,c,d]

[a] Department of Electrical Engineering, The University of Suwon, San 2-2 Wau-ri, Bongdam-eup, Hwaseong-si, Gyeonggi-do, 445-743, South Korea
[b] Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2G6, Canada
[c] Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland
[d] Department of Electrical and Computer Engineering, Faculty of Engineering, King Abdulaziz University, Jeddah, 21589, Saudi Arabia

A R T I C L E I N F O

Article history:
Received 25 August 2017
Received in revised form 19 March 2018
Accepted 27 March 2018
Available online 9 April 2018

Keywords:
Context space
Fuzzy C-Means clustering
Fuzzy partition
Least square estimation

A B S T R A C T

In this study, we introduce a novel category of double fuzzy clustering-driven context neural networks (DFCCNNs). The study is focused on the development of advanced design methodologies for redesigning the structure of conventional fuzzy clustering-based neural networks. The conventional fuzzy clustering-based neural networks typically focus on dividing the input space into several local spaces (implied by clusters). In contrast, the proposed DFCCNNs take into account two distinct local spaces called context and cluster spaces, respectively. Cluster space refers to the local space positioned in the input space whereas context space concerns a local space formed in the output space. Through partitioning the output space into several local spaces, each context space is used as the desired (target) local output to construct local models. To complete this, the proposed network includes a new context layer for reasoning about context space in the output space. In this sense, Fuzzy C-Means (FCM) clustering is useful to form local spaces in both input and output spaces. The first one is used in order to form clusters and train weights positioned between the input and hidden layer, whereas the other one is applied to the output space to form context spaces. The key features of the proposed DFCCNNs can be enumerated as follows: (i) the parameters between the input layer and hidden layer are built through FCM clustering. The connections (weights) are specified as constant terms being in fact the centers of the clusters. The membership functions (represented through the partition matrix) produced by the FCM are used as activation functions located at the hidden layer of the "conventional" neural networks. (ii) Following the hidden layer, a context layer is formed to approximate the context space of the output variable and each node in context layer means individual local model. The outputs of the context layer are specified as a combination of both weights formed as linear function and the outputs of the hidden layer. The weights are updated using the least square estimation (LSE)-based method. (iii) At the output layer, the outputs of context layer are decoded to produce the corresponding numeric output. At this time, the weighted average is used and the weights are also adjusted with the use of the LSE scheme. From the viewpoint of performance improvement, the proposed design methodologies are discussed and experimented with the aid of benchmark machine learning datasets. Through the experiments, it is shown that the generalization abilities of the proposed DFCCNNs are better than those of the conventional FCNNs reported in the literature.

1. Introduction

With the development of neural network and fuzzy logic, a number of neuro fuzzy systems emerged and were applied to many applications in different research and problem domains (Kar, Das, & Ghosh, 2014; Karakuzu, Karakaya, & Çavuşlu, 2016; Mohammed & Lim, 2017; Wu & Zeng, 2016). In early 2000, hybrid neuro fuzzy model based on Takagi–Sugeno–Kang (TSK) fuzzy models have been proposed by Oh and Pedrycz. The centers of membership functions located in the premise part of the model are determined through clustering and then a location of these centers is adjusted through evolutionary computing such as genetic algorithm (GA) and particle swarm optimization (PSO) (Huang & Ding, 2011; Huang, Wang, & Liao, 2016). Moreover, "conventional" clustering-based neural networks are regarded as hybrid models combining radial basis function neural networks (RBFNNs) with clustering algorithm where clustering serves as a learning algorithm for determining the parameters of the activation functions (receptive fields) (Brromhead & Lowe, 1988; Oh, Kim, Pedrycz, & Joo, 2012; Oh, Pedrycz, & Park, 2002, 2003; Roh, Oh, & Pedrycz, 2011; Schwenker, Kestler, & Palm, 2001). At the beginning, *k*-means and Fuzzy C-Means clustering algorithms were used to determine the parameters of Gaussian function forming hidden layer of RBFNNs. The cen-

* Corresponding author.
E-mail address: ohsk@suwon.ac.kr (S.-K. Oh).

https://doi.org/10.1016/j.neunet.2018.03.018
0893-6080/

ters and widths of Gaussian functions are equivalent to the prototype of cluster and the standard deviation of the input variable. This method focuses on identifying the parameters of RBFNNs without structural proceeding with any modification. Since then, a new way to replace the nodes in hidden layer by clusters has been proposed. As a result, determining types and parameters (center and width) of activation function is not required, and instead the membership grade coming from the partition matrix is directly regarded as the output of the hidden layer (Roh, Joo, Pedrycz, & Oh, 2010; Oh, Kim, Pedrycz, & Part, 2011). As another approach, Conditional Fuzzy C-Means (CFCM) clustering was proposed by Pedrycz in 1996. CFCM clustering takes into account the proximity of conditional variables formed in the feature space and conditional variables exhibit effect on forming clusters (groups) of input variables (Pedrycz, 1996). In addition, a single or double clustering has been applied to many models such as evolving fuzzy neural networks (EFuNNs), dynamic evolving neural–fuzzy inference system (DENFIS) and dynamic evolving spiking Neural Networks (DeSNN) (Kasabov, 1998a, 2001; Schliebs & Kasabov, 2013; Soltic & Kasabov, 2010). The structure of DENFIS is similar to the TSK fuzzy model, however DENFIS model exhibits more advanced features of adaptivity from the perspective of underlying structure and learning mechanism in order to form fuzzy partition in input space. Evolving clustering method (ECM) is used to implement a scatter partitioning of the input space to support a formation of the fuzzy rule. DENFIS can be constructed by a variety of clustering technique including both its online and offline versions (Kasabov, 1998a,b; Kasabov & Song, 2002). The DeSNN uses both rank-order learning and dynamic synapses in fast and online mode. The DeSNN uses both rank-order learning and dynamic synapses in fast and online mode. The DeSNN also evolves its structure and functionality in an online manner based on an incoming data. For every new input vector, a new output neuron is dynamically allocated and connected to the input neurons. The weights of the output neurons represent centers of the clusters located in the problem space and can be directly associated with fuzzy rules (Kasabov, Dhoble, Nuntalid, & Indiveri, 2013).

In this study, double fuzzy clustering-driven context neural networks (DFCCNNs) are proposed with the use of the context space derived from the output space by FCM clustering. In the DFCCNNs, FCM is applied two times for both input and output spaces. First, the input space is divided into several fuzzy local spaces by invoking the FCM algorithm. At the second phase, the output space is also split into several fuzzy local spaces called contexts. In general, the centers of membership (activation) functions were selected as uniform intervals without definite criteria for considering the distribution and characteristic of the given data. That is to say, since the real world problem (data) involves the high complexity such as the high dimension, irregular distribution and hidden nature, the evenly partition-based membership function cannot be suitable as the approach to solve the given problem. However, clustering can offer an alternative solution in order to form variable (adjustable) membership functions focused on inherent aspects of the given problem. Clustering discovers centers of clusters (groups) through iterative adjustments of prototypes and partition matrix. In clustering-based neural networks, in order to enhance the interpretability of the input space, the clustering technique is usually applied to data located in this space. Using regression methods, the mapping between each local space and the desired output is realized (Oh et al., 2012, 2002, 2003; Roh et al., 2011). Since the desired output is applied equally to all local models, the performance or accuracy of all local models is similar one another. An effect of merging local model is low. However, the local models of the proposed DFCCNNs are made through fuzzy regression (inference) between the cluster in input space and the context in the output space. A degree of membership of clusters is used as inputs in granular (i.e. fuzzy) space whereas a degree of membership of context is used as a desired output (i.e. original output) to each local model. Namely, the values of desired output are differently given to every local models. Hence, the local models of the proposed networks are constructed more independently than the case in the conventional clustering-based neural networks. The final outputs of the DFCCNNs are merged through a weighted average of the local models formed in the output layer.

The paper is structured as follows. In Section 2, we present the architecture of the double fuzzy clustering-driven context neural network and discuss the underlying learning procedure realized by means of the FCM and LSE in Section 3. Section 4 covers an overall design methodology while experimental results are presented in Section 5. Conclusions are drawn in Section 6.

2. Architecture of double fuzzy clustering-driven context neural networks

2.1. Basic concept of the double fuzzy clustering-driven context neural networks

About a decade ago, conditional fuzzy c-means (CFCM) clustering was proposed by Pedrycz. A Conditional variable in CFCM serves as a navigation mechanism for forming fuzzy local spaces positioned in input space. Since then, linguistic model (LM) and granular neural network (GNN) started to use the conditional FCM clustering using the membership grades of the output as a certain conditional variable (Pedrycz & Kwak, 2006; Pedrycz, Park, & Oh, 2008; Park, Pedrycz, & Oh, 2009). The "conventional" models regard a set of the membership grades (values) of the output variable as context and use context-based FCM clustering to form clusters of the input space based on each context space as shown in Fig. 1(a). Fig. 1(a) shows the process of generating new clusters in the input space by the guide of the context of output space, and conventional models such as linguistic model and granular neural network use the clusters as the nodes in the hidden layer (Pedrycz, 1996).

In this study, we use a set of contexts derived from output as the granular (fuzzy) output to construct local model and to train connections (weights) corresponding to local model. Unlike the conventional model, we do not use the conditional FCM to form clusters in the input space. Instead, we use FCM clustering in both input and output spaces to form cluster and context separately, and the elements belonging to clusters or contexts come with some membership grades determined by the corresponding entries of the partition matrix. Hence, the size or structure of the space is similar to the one being encountered in the conventional model such as LM and GNN, but the role of the context is different. As shown in Fig. 1(b), once both spaces have been formed by the FCM clustering, the context of the proposed DFCCNNs is inferred through the clusters located in the input space and the connections (shown as arrows in this figure) corresponding to clusters. Then, the numeric value in the output space is calculated through decoding or aggregation the inferred contexts. To do this, fuzzy inference scheme based on centroid defuzzifier is applied.

The difference in which the contexts between the convention model and the proposed DFCCNNs model are used can be summarized in Table 1.

Fig. 2 shows the algorithmic details of the proposed DFCCNNs. First, input space ($\boldsymbol{X}$) is transformed into fuzzy local spaces ($\boldsymbol{U}$) by running the FCM clustering, which means that n-dimensional input variables are converted c-dimensional fuzzy granular variables. Each component of granular variables means a local space. Next, the approximation space ($\hat{\boldsymbol{T}}$) of context is inferred by a linear combination of the fuzzy local spaces ($\boldsymbol{X}$) and weights ($\boldsymbol{W}$). At last, the final output ($\hat{\boldsymbol{Y}}$) is calculated through the fuzzy inference of the estimated (inferred) context ($\hat{\boldsymbol{T}}$) and another connections ($\boldsymbol{H}$).