www.ijecs.in
International Journal Of Engineering And Computer Science ISSN: 2319-7242
Volume 5 Issue 6 June 2016, Page No. 16753-16755

Multitask TSK Fuzzy System Modeling by Mining Intertask Common Hidden Structure

Sonawane Gokul R.[1] Gaikwad Vikas R.[2] Vaidya Nikhil R.[3] Sudake Shyam S.[4] Prof. Rokade P.P

IT Department,
SND college of Engineering, Yeola,Pune University.
gokulsonawane9@gmail.com
IT Department,
SND college of Engineering, Yeola,Pune University.
vikasgaikwad201415@gmail.com
IT Department,
SND college of Engineering, Yeola,Pune University.
nikhilvaidya.1992@gmail.com
It Department,
SND college of Engineering, Yeola,Pune University.
shyamsudke111@gmail.com

Head, Dept. of IT Engineering SND College of Engineering, Yeola
@gmail.com

***Abstract*— The classical fuzzy system model method kindly assumed data which is generated from single task. This data can be acquired from the perspective of multiple task the modeling has on intrinsic inconsistency. In this project , a multiple fuzzy system modeling method by mining interact common hidden structure is propose to overcome the weakness of classical TSK- based fuzzy modeling method system for multitask learning. When the classical fuzzy modeling method are applied to multitask datasets, they usually focus on the task independence information and the ignore the correlation between different task. Here we mine the common hidden structure among multiple tasks to realize multitask TSK fuzzy system learning it makes good used of the independence information of each task and correlation information captured by common hidden structure among all tasks as well. Thus, the proposed learning algorithm can effectively improve both the generalization and fitting performance of the learned fuzzy system for each task . Our experiment result demonstrate at the proposed MTCS-TSK_FS has better modeling performance and adaptability than the existing TSK based fuzzy modeling method on multitask datasets. Learning multiple tasks across different datasets is a challenging problem since the feature space may not be the same for different tasks. The data can be any type or the datasets. The data can be any type or the datasets are any type like text datasets.**

Index Terms—Common hidden structure, fuzzy modeling,multitask learning, Takagi-Sugeno-Kang (TSK) fuzzy systems.

Introduction

The learning multiple related task simultaneously has better performance than learning these tasks independently. Focused on multitask learning, the main goal of multitask learning is to improve the generalization performance of learners by leveraging the domain-specific information contained in the related tasks. The way to reach the goal is learning multiple-related tasks simultaneously by using a common representation. In fact, the training signals for extra tasks serve as an inductive bias which is used to learn multiple complex tasks together. Theoretical studies on multitask learning have been actively performed in the following three areas: multitask classification learning , multitask clustering and multitask regression learning.

It has been shown in those studies that, when there are relations between multiple tasks to learn, it is beneficial to learn them simultaneously instead of learning each task independently. Although those studies have indicated the significance of multitask learning and demonstrated certain effectiveness in different real-world applications, the current multitask learning methods are still very limited and cannot keep up with the real-world requirements, particularly in fuzzy modeling.

Thus, this project focuses on multitask fuzzy modeling. With rapid development of data collection technologies, the regression datasets are collected with different formats. Thus, the proposed multitask regression learning method makes good use of both task independent information and inter-task hidden correlation information of a given dataset.

I. LITERATURE SURVEY

ARECENT principle tells us that multitask learning or learning multiple related tasks simultaneously has better performance than learning these tasks independently [1]–[2]. Focused on multitask learning, the principal goal of multitask learning is to improve the generalization performance of learnerg by leveraging the domain-specific information contained in the related tasks [2]. One way to reach the goal is learning multiple-related tasks simultaneously while using a common representation. In fact, the training signals for extra tasks serve as an inductive bias which is helpful to learn multiple complex tasks together [3]. Empirical and theoretical studies on multitask learning have been actively performed in the following three areas: multitask classification learning [2]–[5}, multitask clustering [2]–[4], and multitask regression learning [2]–[5]. It has been shown in those studies that, when there are relations between multiple tasks to learn, it is beneficial to learn them simultaneously instead of learning each task independently. Although those studies have indicated the significance of multitask learning and demonstrated certain effectiveness in different real-world applications, the current multitask learning methods are still very limited and cannot keep up with the real-world requirements, particularly in fuzzy modeling. Thus

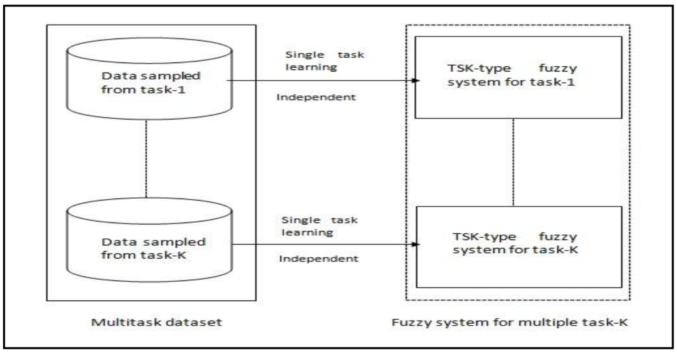

Fig. 1. Classicale Fuzzy Model.

In classical multitask fuzzy modeling method to learn multiple task we have to build individual fuzzy model for each task, There is lot of time wasted for build fuzzy model for each task to save the time for individual fuzzy model we can build multitask fuzzy model.

II. PROPOSED SYSTEM

This project focuses on multitask fuzzy modeling[10]. With the rapid development of data collection technologies, the regression datasets are collected with different formats. While each resulted dataset can be modeled individually to preserve the independence of each task ,

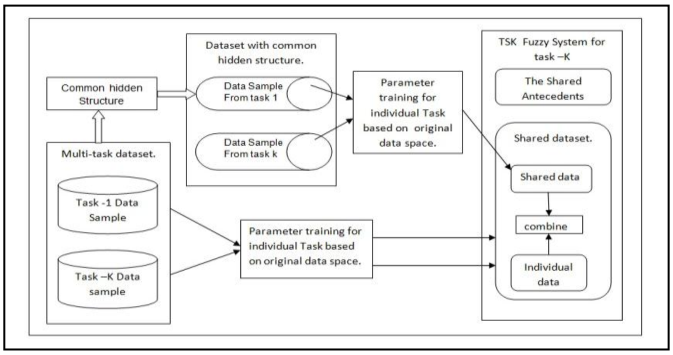

Fig. 2. Architecture Of Proposed System.

the inter-task hidden correlation is virtually lost. In order to solve this problem, this project proposes a novel modeling technique which can effectively take into consideration of the hidden correlation information between datasets. Thus, the proposed multitask regression learning method makes good use of both task independent information and inter-task hidden correlation information of a given dataset.

1. Vector Model :-

vector model is the algebraic model used for representing text documents (and any objects, in general) as vectors of identifiers, such as, for example, index terms.
Documents and queries are represented as vectors.

$$d_j = (w_{1,j}, w_{2,j}, \dots, w_{t,j})$$
$$q = (w_{1,q}, w_{2,q}, \dots, w_{n,q})$$

Each dimension corresponds to a separate term. If a term occurs into the document, its value in the vector is non-zero. Several different ways of computing these values, also known as weights, have been developed. One of the best known schemes is tf-idf weighting

2. TF–IDF Weights :-

In the classic vector space model proposed by Salton, Wong and Yang the term-specific weights in the document vectors are products of local and global parameters. The model is known as term frequency-inverse document frequency model[7]. The weight vector for document *d* is $\mathbf{v}_d = [w_{1,d}, w_{2,d}, \dots, w_{N,d}]^T$,
where

$$w_{t,d} = \mathrm{tf}_{t,d} \cdot \log \frac{|D|}{|\{d' \in D \,|\, t \in d'\}|}$$

- $\mathrm{tf}_{t,d}$ is term frequency of term *t* in document *d* (a local parameter)
- $\log \frac{|D|}{|\{d' \in D \,|\, t \in d'\}|}$

is inverse document frequency (a global parameter). $|D|$ is the total number of documents in the document set; $|\{d' \in D \,|\, t \in d'\}|$ is the number of documents containing the term *t*.
Using the cosine the similarity between document d_j and query *q* can be calculated as:

$$\mathrm{sim}(d_j, q) = \frac{\mathbf{d_j} \cdot \mathbf{q}}{\|\mathbf{d_j}\| \, \|\mathbf{q}\|} = \frac{\sum_{i=1}^{N} w_{i,j} w_{i,q}}{\sqrt{\sum_{i=1}^{N} w_{i,j}^2} \sqrt{\sum_{i=1}^{N} w_{i,q}^2}}$$

3. Support Vector Machine" (SVM) :-

"Support Vector Machine" (SVM) is a supervised machine learning algorithm which can be used for both classification or regression challenges[6]. However, it is mostly used for the classification problems. In this algorithm, we plot each data item as a point in n-dimensional space with the value of each feature being the value of a particular coordinate. Then, we perform classification by finding the hyper-plane that differentiate the two classes very well (look at the below snapshot).

4. SVMs results :-

There is much more in the field of SVMs than we could Cover here, including [8]:

- Regression, clustering, semi-supervised learning and other domains. Lots of other kernels, e.g. string kernels to handle text.
- Lots of research in modifications, e.g. to improve generalization ability, or tailoring to a particular task.. Lots of research in speeding up training

5. Common Hidden Structure :-

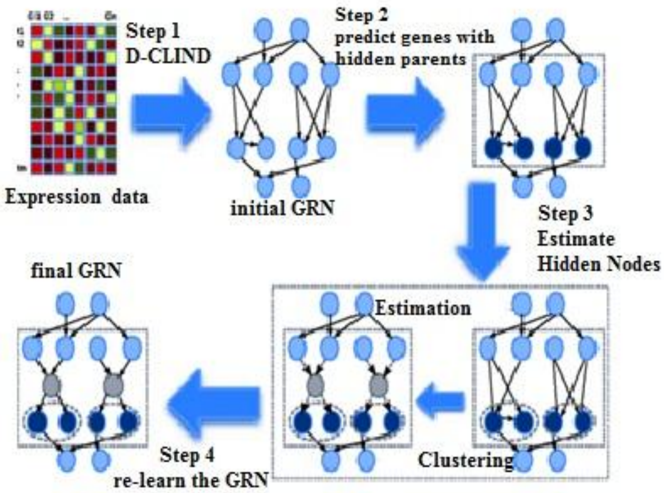

Fig. 3. Steps To Finding Common Hidden Structure.

Overall flow And steps of the proposed Algorithm :-

1 Infer an initial GRN,

2 Identify the genes with hidden common cause,

3 Estimate the hidden common cause(s), which involves clustering and EM,

4) Re-learn the GRN after estimation of the hidden common cause(s)

III. ALGORITHMIC STRATERGIES TO BE USED:-

SVM Algorithm [9]:-

Require: A linearly separable set S, learning rate $\eta \in \Re^+$

$w_0 = 0; b_0 = 0; k = 0;$

$R = \max_{1 \le i \le l} ||x_i||$

while at least one mistake is made in the **for** loop **do**

  for $i = 1, \ldots, l$ **do**

    if $y_i(< w_k, x_i > +b_k) \le 0$ **then**

      $w_{k+1} = w_k + \eta y_i x_i$

      $b_{k+1} = b_k + \eta y_i R^2$ (updating bias[1])

      $k = k + 1$

    end if

  end for

end while

Return w_k, b_k, where k is the number of mistakes

IV. CONCLUSION FUTURE SCOPE

In this project, a multitask fuzzy system modeling method by mining interact common hidden structure is propose to overcome the weakness of classical TSK-based fuzzy modeling method system for multitask learning. When the classical fuzzy modeling method are applied to multitask datasets, they usually focus on the task independence information and the ignore the correlation between different task.

Here we mine the common hidden structure among multiple tasks to realize multitask TSK fuzzy system learning. It makes good used of the independence information of each task and the correlation information captured by the common hidden structure among all tasks as well. Thus, the proposed learning algorithm can effectively improve both the generalization and fitting performance of the learned fuzzy system for each task. Our experiment result demonstrate at the proposed MTCSTSK-FS has better modeling performance and adaptability than the existing TSK based fuzzy modeling methods on multitask datasets.

REFERENCES

1 R. Caruana, “Multitask learning,” Mach. Learn., vol. 28, no. 1, pp. 41–75, 1997.

2 S. Sun, “Multitask learning for EEG-based biometrics,” in Proc. 19Th Int. Conf. Pattern Recognit., Tampa, FL, USA, 2008, pp. 1–4.

3 X. T. Yuan and S. Yan, “Visual classification with multi-task joint sparse representation,” in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., San Francisco, CA, USA, 2010, pp. 3493–3500.

4 T. Jebara, “Multi-task feature and kernel selection for SVMs,” in Proc. 21st Int. Conf. Mach. Learn., Banff, AB, Canada, Jul. 2004.

5 S. Kong and D. Wang, “A multi-task learning strategy for unsupervised clustering via explicitly separating the commonality,” in Proc. ICPR, Tsukuba, Japan, 2012, pp. 771–774.

6 J. Weston and C. Watkins, “Support vector machines for multi-class pattern recognition,” in *Proc. 7th Eur. Symp. Artif. Neural Netw.*, 1999, pp. 219–224.

7 Ben Hachey, “Multi-document summarization using generic relation extraction**”,** Proceedings of the ACL Conference on Empirical Methods in Natural Language Processing: Volume 1, 420-429, 2009.

8 Y. Chen and J. Z. Wang, “Support vector learning for fuzzy rule-based classification systems,” *IEEE Trans. Fuzzy Syst.*, vol. 11, no. 6, pp. 716–728, Dec. 2003.

9 T. Heskes. Solving a huge number of similar tasks: a combination of multi-task learning and hierarchical Bayesian modeling. In *ICML*, pages 233–241, 1998.

10 S. L. Chiu, “Fuzzy model identification based on cluster estimation,” *J.Intell. Fuzzy Syst.*, vol. 2, pp. 267–278, 1994.