Fuzzy Signature Neural Networks
for Classification: Optimising the Structure

Tom Gedeon, Xuanying Zhu, Kun He, and Leana Copeland

Research School of Computer Science, College of Engineering and Computer Science,
The Australian National University, Canberra, ACT 0200, Australia
{tom.gedeon,u5251881,u5058161,leana.copeland }@anu.edu.au

Abstract. We construct fuzzy signature neural networks where fuzzy signatures replace hidden neurons in a neural network similar to a radial basis function neural network. We investigated the properties of a naïve and a principled approach to fuzzy signature construction. The naïve approach provides very good results on benchmark datasets, but is outperformed by the principled approach when we approximate the noisy nature of real world datasets by randomly eliminating 20% of the data. The major benefit of the principled approach is to substantially improve robustness of the fuzzy signature neural networks we produce.

Keywords: fuzzy signature, neural network, RBF, clustering, salary selection, diabetes.

1   Introduction

Given k inputs, and at most T linguistic terms per dimension of X for the α-cover, the number of fuzzy rules covering X at least to α is $|R| = \mathrm{O}(T^k)$ which is very high, unless k is very small. The exponential explosion in rules is a major problem hindering the application of fuzzy techniques beyond control systems. We have previously developed and adapted techniques which partially address decreasing T, and of k (by rule interpolation and hierarchical rule bases respectively [1-3]. Encouraging results of our previous research give further motivation to continue along these lines, and to develop further new techniques that will be suitable for the solution of even harder problems.

We have developed two kinds of approaches to the exponential explosion. First, sparse hierarchical fuzzy systems reduce both T and k simultaneously by finding (top-down) sub-spaces in the data, which allows some dimensions and rules to be ignored. The use of sparse rule base allows proper fuzzy reasoning even if the rule set contains "gaps" [4]. For various technical reasons, eliminating dimensions and rules at the same time gives the best results from the produced rule bases and is at the same time the most efficient [5].

The second is fuzzy signatures – constructing characteristic fuzzy structures, modelling the complex structure of the data points (bottom up) in a hierarchical manner.

Fuzzy signatures result in a much reduced order of complexity, at the cost of more complex aggregation techniques.

General fuzzy signatures are vector valued fuzzy sets, where each vector component can be a further vector valued fuzzy set, and use aggregations to propagate the fuzzy values from low levels to high levels in the structures, resulting in effective and efficient fuzzy inference. Such aggregations encompass the simple classic fuzzy conjunction and union operations, but still maintain transparency of fuzzy reasoning.

A Fuzzy Signature can be graphically represented by two structures, hierarchical structure, which is a vector, and a tree. Figure 1 below illustrates these two Fuzzy Signature structures for a SARS patient [6], conveying the same information.
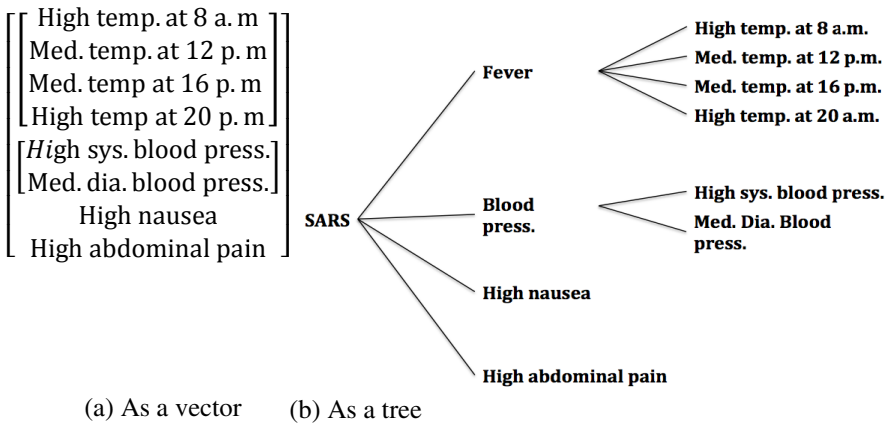

(a) As a vector   (b) As a tree

Fig. 1. Example of two Fuzzy Signature structures for a SARS patient

2   Fuzzy Signature Neural Networks

A fuzzy signature neural network (FZSIGNN) is similar to a radial basis function (RBF) neural network, where the RBF neurons are replaced by fuzzy signature neurons, where each neuron is a fuzzy signature. RBF neural networks are faster to train than back-propagation neural networks of a similar topology, as only the output layer of weights is trained.

We would expect a fuzzy version to be faster, since we would use fuzzy membership regions which are rectilinear as opposed to the (usually Euclidean) distance calculations required by RBF networks.

Each hidden neuron has a particular fuzzy signature associated with it, and outputs the similarity between input vector and fuzzy signature. This means that the inputs to the hidden layer are unmodified values read directly from input data. Hence, the weights connected between input layers and hidden layers are constants. Therefore, all the learning is done via the weights connected between hidden layer and output layer. Thus, the time to train these neural networks should be reduced due to decreasing size of weight matrix, and eliminating the need to propagate errors to slowly train layers of weights further from the output. The construction of the fuzzy

signatures may take some time, which is investigated in this paper, in comparing a naïve but fast approach with a principled but slower approach.
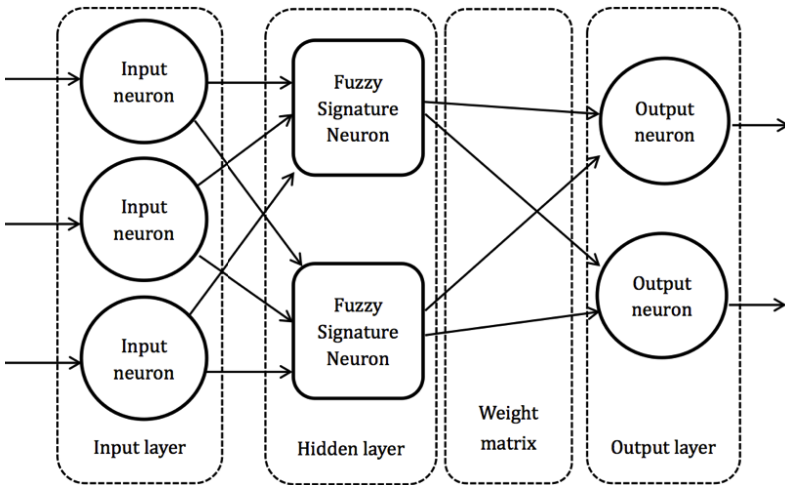
Fig. 2. Fuzzy Signature Neural Network structure

3   Fuzzy Signature Construction

In both of our techniques we use hierarchical agglomerative clustering since it does not require users to specify the number of clusters and contains more information than flat clustering [7].
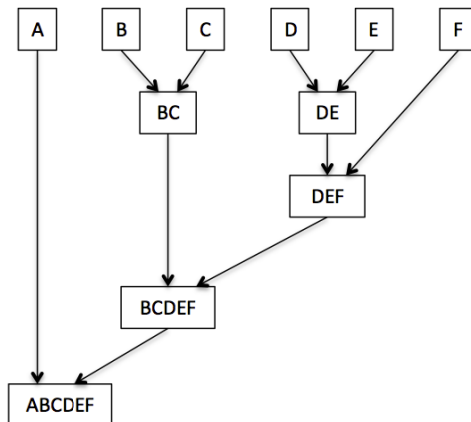
Fig. 3. Example of Agglomerative hierarchical clustering

We use various heuristics to restrict the number of clusters to 5, and use the cluster centroids to generate fuzzy signatures. In the naïve approach, *2* to *n-1* input variables are chosen at random with aggregation function chosen from the set { max, min, ave } by calculation of the max, min and ave of the membership values and selecting as aggregation function the one with the smallest standard deviation [8]. In our principled approach, we use genetic programming to select both the number of variables and the aggregation functions used.

Table 1. Genetic programming settings for constructing the fuzzy signatures

Parameters	Values
Number of generations in GPLAB	10
Number of individuals in GPLAB	50
Number of fuzzy signature neurons	5
Training epoch	100
Percentage of training data	80%
Percentage of testing data	20%
Percentage of missing values	0% / 20%

4   Results

Four datasets are used for evaluation, two of which are Salary problems from Gedeon [9], one is Salary problem from Mendis [10], and the rest are Cancer and Diabetes problems from University of California Irvine (UCI) [11]. Table 2 below shows the general information about these datasets.

4.1   Dataset Properties

Table 2. Datasets used

Dataset	Number of input attributes	Number of output columns	Number of observations
Diabetes	8	2	768
High Salary (Gedeon)	3	2	200
Medium Salary (Gedeon)	3	2	200
Low Salary (Mendis)	3	2	135

4.2   Clean Data

We first used our two approaches on the clean datasets with no changes.

Table 3. Using principled approach to construct fuzzy signatures

Dataset	Testing		
	Mean (%)	StDv	MSE
Diabetes	68.5	0.0184	0.2406
High Salary (Gedeon)	81.5	0.0141	0.1903
Medium Salary (Gedeon)	83.0	0.0283	0.1423
Low Salary (Mendis)	80.1	0.0406	0.1607
Average	**78.3**	**0.0254**	**0.1835**

We provide average prediction performance results only for rough comparison between our two techniques, as such averages do not make sense in general.

Table 4. Using naïve approach to construct fuzzy signatures

Dataset	Testing		
	Mean (%)	StDv	MSE
Diabetes	68.2	0.0359	0.1987
High Salary (Gedeon)	78.0	0.0622	0.1160
Medium Salary (Gedeon)	88.5	0.0379	0.0876
Low Salary (Mendis)	68.9	0.1425	0.1934
Average	**75.9**	**0.0696**	**0.1489**

We can see that the naïve approach is worse on average, though on the simplest dataset (medium salary) it performs the best. On the hardest dataset (low salary) it performs the worst.

4.3   Damaged Data

We then used our two approaches on the datasets with 20% of the data omitted at random.

Table 5. Principled approach on damaged data

Dataset	Testing		
	Mean (%)	StDv	MSE
Diabetes	65.1	0.0147	0.2438
High Salary (Gedeon)	83.0	0.0694	0.1428
Medium Salary (Gedeon)	87.5	0.0530	0.1077
Low Salary (Mendis)	74.8	0.1154	0.1805
Average	**77.6**	**0.0631**	**0.1687**

On the hardest dataset (low salary) and on the diabetes dataset, the performance has decreased but on the other two dataset including the easiest dataset, the

performance has actually increased. We believe this is due to elimination of some outliers from a dataset which otherwise has redundant data. We have seen such results in our previous work in heuristic pattern reduction and bimodal distribution removal. Overall, the principled approach is quite robust to the damaged data.

Table 6. Naïve approach on damaged data

Dataset	Testing		
	Mean (%)	StDv	MSE
Diabetes	64.6	0.0461	0.2180
High Salary (Gedeon)	79.5	0.0209	0.1485
Medium Salary (Gedeon)	77.5	0.0467	0.1614
Low Salary (Mendis)	55.6	0.0642	0.2388
Average	**69.3**	**0.0445**	**0.1917**

The naïve approach has reduced performance quite significantly overall, but has retained its performance on one of the datasets.

5   Conclusion and Future Work

We have shown our proposed approach to produce fuzzy signature neural networks, with both a naïve and a principled approach to construction of the fuzzy signature neurons.

The principled approach is more expensive computationally but is robust in the face of significant damage to the datasets of 20% deletion at random.

On the other hand, the naïve approach is computationally cheap, and performed well on at least one dataset even with damage.

Our future work will include identifying the applicability conditions where the naïve approach can be expected to perform well, and further improvements of our principled approach including particularly reduction of the computational complexity – it may be possible in certain circumstances which we intend to delineate to replace the global search of the evolutionary approach we used with a local gradient descent search.

References

1. Tikk, D., Baranyi, P., Gedeon, T.D., Muresan, L.: Generalization of the rule interpolation method resulting always in acceptable conclusion. Tatra Mountains Math. Publ. 21, 73–91 (2001)

2. Mendis, B.S.U., Gedeon, T.D., Kóczy, L.T.: Investigation of aggregation in fuzzy signatures. In: 3rd International Conference on Computational Intelligence, Robotics and Autonomous Systems, vol. 406 (2005)
3. Mendis, B.S.U., Gedeon, T.D., Botzheim, J., Kóczy, L.T.: Generalised weighted relevance aggregation operators for hierarchical fuzzy signatures. In: 2006 International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, pp. 198. IEEE (2006)
4. Tikk, D., Biró, G., Gedeon, T.D., Kóczy, L.T., Yang, J.D.: Improvements and critique on Sugeno's and Yasukawa's qualitative modeling. IEEE Transactions on Fuzzy Systems 10(5), 596–606 (2002)
5. Koczy, L.T., Hirota, K., Gedeon, T.D.: Fuzzy rule interpolation by the conservation of relative fuzziness. International Journal of Advanced Computational Intelligence 4(1), 95–101 (2000)
6. Wong, K.W., Gedeon, T., Kóczy, L.F.: Construction of fuzzy signature from data: An example of SARS pre-clinical diagnosis system. In. In: Proceedings of the IEEE International Conference on Fuzzy Systems, vol. 3, pp. 1649–1654. IEEE (2004)
7. Tevor, H., Robert, T., Jerome, F.: The Elements of Statistical Learning, 2nd edn., pp. 520–528. Springer, New York (2009)
8. Mendis, B.S.U., Gedeon, T.D.: Complex Structured Decision Making Model: A hierarchical frame work for complex structured data. Information Sciences 194, 85–106 (2011)
9. Gedeon, T., Wong, K., Tikk, D.: Constructing hierarchical fuzzy rule bases for classification. In: Proceedings of IEEE International Conference on Fuzzy Systems, pp. 1388–1391 (2001)
10. Mendis, B.S.U.: Fuzzy Signatures: Hierarchical Fuzzy Systems and Applications, PhD thesis, Department of Computer Science, The Australian National University (March 2008)
11. Bache, K., Lichman, M.: The UCI Machine Learning Repository. Center for Machine Learning and Intelligent Systems, University of California, Irvine (1987), `http://archive.ics.uci.edu/ml/datasets.html` (viewed September 15, 2013)