Rule extraction from spatial data using local learning techniques

***Brendon J. Woodford*[1]**

[1]Department of Information Science
University of Otago. Dunedin, New Zealand
Phone: +64 3 479-5397 Fax: +64 3 479-8311
Email: bjwoodford@infoscience.otago.ac.nz

Presented at SIRC 2005 - The 17th Annual Colloquium of the Spatial Information Research Centre
University of Otago, Dunedin, New Zealand
November 24th-25th 2005

ABSTRACT

We are now in the fourth decade where techniques such as fuzzy systems, statistics, neural networks and machine learning have all been developed and more recently applied for the purpose of spatial data mining. However these methods act as global learning models and subsequently may not be able to learn the subtle nature of these types of data sets. Local learning models such as the Support Vector Machine (SVM) and a more recent method such as that proposed by (Gilardi 2002) address the problem of global versus local learning but fail to offer many solutions as to what underlying patterns may exist within the data set in order to better understand the data set. In this paper we propose the Evolving Fuzzy Neural Network (EFuNN) as a model for a local learning mechanism for the purpose of predicting rainfall within a region of Switzerland and also use this model to generate rules and then to visualise these rules that may help to describe any patterns that may exist within the data set.

Keywords and phrases: local learning, similarity metrics, machine learning, fuzzy systems, neurocomputing, rule extraction

1 INTRODUCTION

The prevalence and diversity of spatial data has seen many different machine learning techniques applied for the purpose of spatial data mining (Chawla, Shekhar, Wu & Ozesmi 2001). These methods date back to over four decades ago such as fuzzy logic (Zadeh 1965), statistics (Sammon 1969), Neural Networks (NN) (Rumelhart, Hinton & Williams 1986), and machine learning (Mitchell 1997) and these methods can be characterised by the fact they are global learning mechanisms.

However spatial data is not a normal “beast” and therefore requires a different approach for the purpose of creating better classifiers or predictors. Recently (Gilardi 2002) demonstrated the effectiveness of applying machine learning models that applied local learning algorithms. Here, the objective was to apply local learning address the problem of identifying clusters where it may be difficult to easily separate different regions within the data set.

Although this local-learning model was proved to be effective, it did not explain the nature of the clusters that were identified through the process of local learning and these “nuggets” may have gone some way as to providing some more insight into the nature of the data set.

In this paper we show how applying the EFUNN to the same data set not only can we employ a local learning model on such spatial data but also extract rules from the model in order to help reveal and visualise any patterns or clusters that may exist within the data set so that more refined machine learning models can be created with better generalisation performance.

2 THE MACHINE LEARNING METHODS

2.1 Support Vector Machine

The Support Vector Machine (SVM) (Cortes & Vapnik 1995) is an example of a kernel-machine learning method that where it maps the input vectors of the data set into a higher dimensional space. From this higher dimensional space a linear decision surface is constructed that exhibits special properties that ensure high generalisation ability of the network. Although initially used for classification problems the SVM has now been modified so that it can be applied to regression problems through Support Vector Regression (SVR) (Smola & Schölkopf 1998).

Both SVM and SVR are global learning mechanisms so the approach taken by (Gilardi 2002) was to apply a local learning method based on SVR for the purpose of investigating spatial data sets.

2.2 The EFuNN

An alternative approach would be to use the Evolving Fuzzy Neural Network (EFuNN). This model is a natural extension of the Fuzzy Neural Network (FuNN) (Kasabov, Kim, Watts & Gray 1996) structure whose learning mechanism evolves its structure according to the Evolving Connectionist Systems (ECOS) principles (Kasabov 1998a, Kasabov 1999a, Kasabov 1999b, Kasabov 1998b). Although it is based on work adapted from (Amari & Kasabov 1997, Carpenter & Grossberg 1991, Kohonen 1990, Kohonen 1997), this architecture also introduces some new NN techniques. For example all nodes in an EFuNN are created during (possibly one-pass learning). For a detailed explanation of the EFuNN evolving algorithm we direct the reader to (Kasabov 1998a, Kasabov 1998b, Kasabov 1998c, Kasabov 1999a, Kasabov 2001).

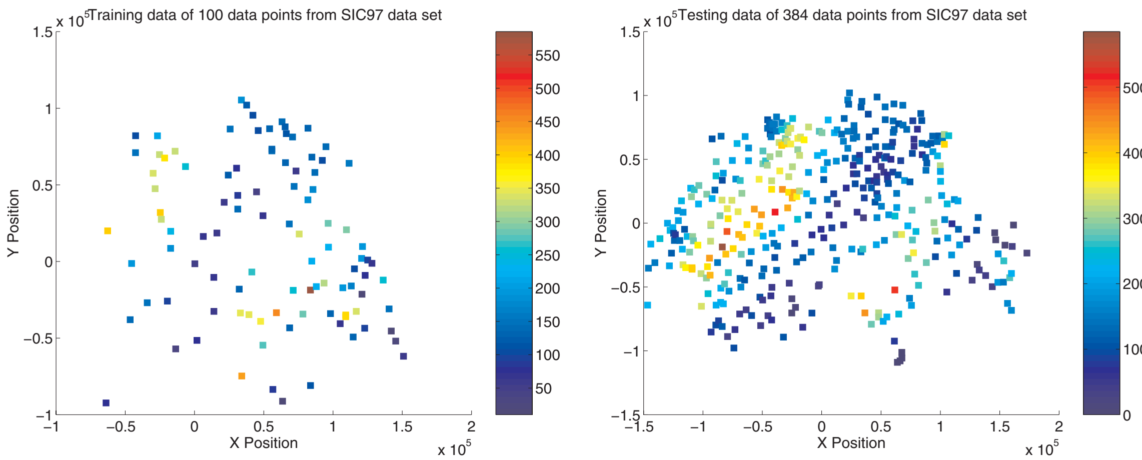
The EFuNN is a good candidate to compare its performance against the local SVM method. The reasons are fourfold:

1. The EFuNN is a local learning model as well.
2. It utilises one-pass real-time learning.
3. It is fast, adaptive and it has been demonstrated that it exhibits good generalisation
4. More importantly, it allows for extraction of rules from the model using an algorithm described in (Kasabov & Woodford 1999) and this characteristic may help to explain why local learning models perform better on such spatial data sets when visualised.

3 THE CASE STUDY

The data set in question was compiled for a special issue of the journal of Geographic Information and Decision Analysis (GIDA) and called the Spatial Interpolation Comparison 97 (SIC 97) (Dubois, Malzekski & DeCort 1998). Here the problem was to estimate the simultaneous local rainfall over a large region of Switzerland when provided with the X-Y coordinates of a digital evaluation system and a measurement of rainfall at each location. There were 484 rows within the data set but interestingly the only 100 measurements were given and the submitted models were asked to predict the rainfall in the other 384 locations.

Figure 1(a) is a plot of the 100 instances of the training data and Figure 1(b) are the 384 data points to be predicted.



(a) Training data

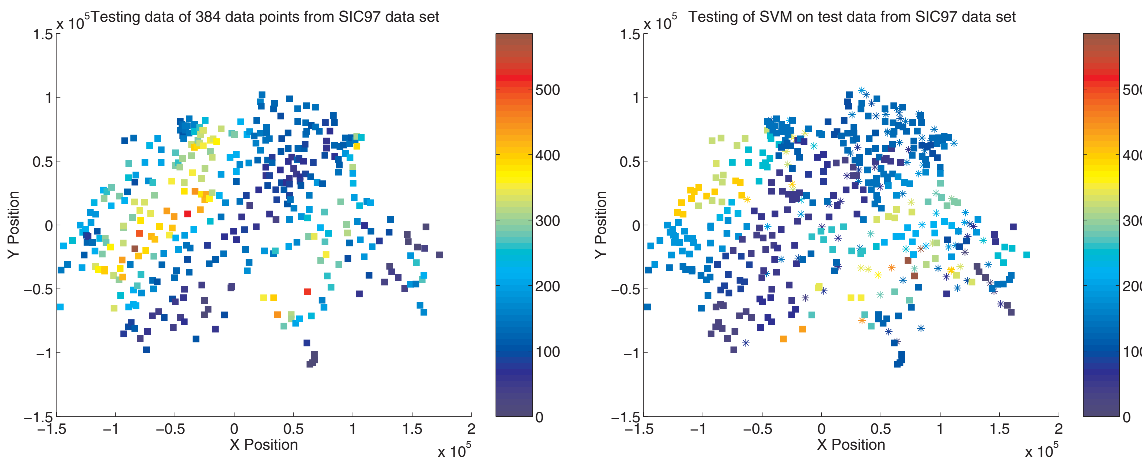
(b) Testing data

Figure 1: *The Spatial Interpolation Comparison 97 data set*

Such a data set presents a challenge to any type of learning model whether it be a global learning or local learning model. Notice that the training data set 1(a) the data points are sparsely located while the testing data set in 1(b) is more densely clustered. We would then expect that the learning model would have to be very good at generalisation to accommodate the unseen data examples.

3.1 Training and testing the SVM

For the version of the SVM we used an implementation from the OSU SVM Classifier Matlab Toolbox (Ma & Ahalt 2001). Using 45 Radial Basis Function (RBF) kernels, the SVM was training on the 100 instances of the training data and then subsequently tested on the test data set. The results of the test data are in Figure 2(a) and the results of testing the SVM on the test data are in Figure 2(b). The asterisks represent the kernel centres of the SVM.



(a) Actual test data from data set

(b) Testing of SVM on test data set

Figure 2: *Local learning results from SVM*

We note that even with local learning, the SVM was not easily able to generalise to the unseen testing data especially on the left hand side of 2(b) where there are no kernel centres present at all.

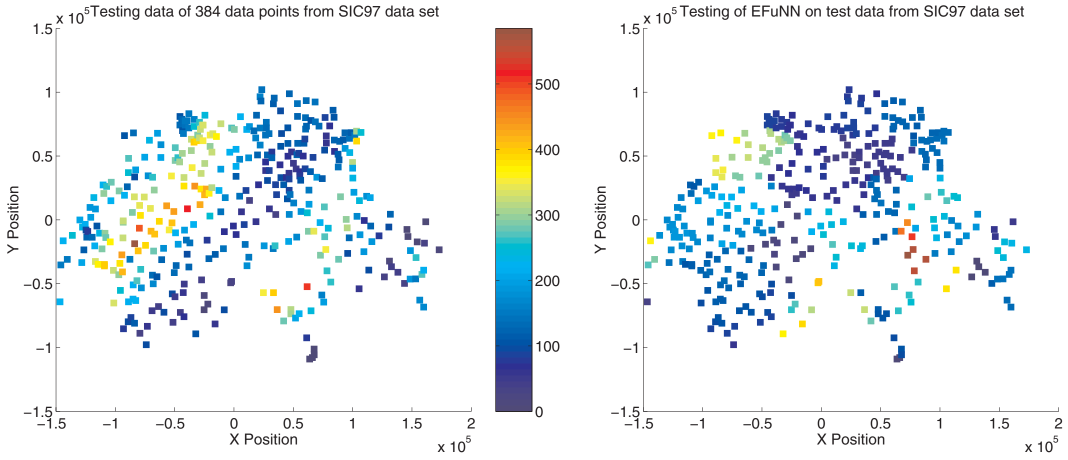
3.2 Training and testing the EFuNN

An EFuNN was set up with an architecture of

- 2(inputs)-5(inputMF)-0(ruleNodes)-5(outputMF)-1(output)

The sensitivity threshold *Sthr*=0.1, error threshold *Errthr*=0.01, and learning rate for both the first and second layer *lr*=0.5. These parameters were selected to create an EFuNN with a large number of rule nodes to accommodate the sparse distributed nature of the training data set.

After training the number of rule nodes generated was 63 and the Root Mean Squared (RMS) error of the tested EFuNN was 0.049. The results of the test data are in Figure 3(a) and the results of testing the EFuNN on the test data are in Figure 3(b).



(a) Actual test data from data set

(b) Testing of SVM on test data set

Figure 3: *Local learning results from EFuNN*

3.3 An analysis of the rules extracted from the EFuNN

To make more sense of what rule node was allocated to each data instance within the training set, rule extraction was carried out on the EFuNN. This produced 63 rules and the visualisation is contained in Figure 4 where each data instances has an accommodated rule node beside it. Here it is noted that there is much more evidence of local learning from the EFuNN as data instances of similar values were accommodated by the same rule.

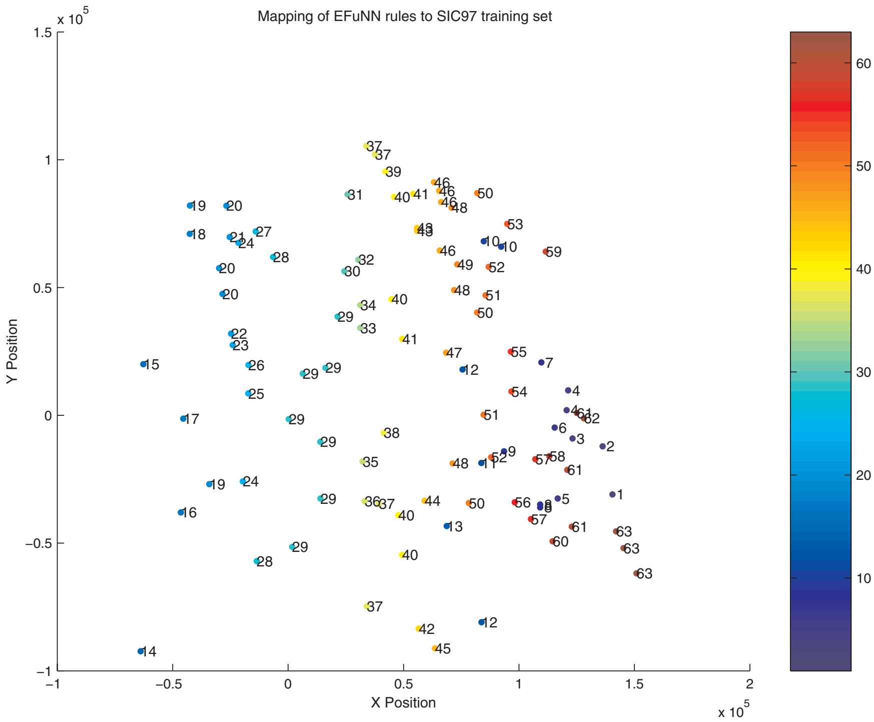


Figure 4: *Rule nodes mapped to training data set*

There is more evidence of this when we investigate the performance of the EFuNN on the test data set using the same data visualisation. Here we can easily see that the clustering in Figure 5 is more apparent and that each rule is accommodating more data examples but in a much larger but still bounded region.

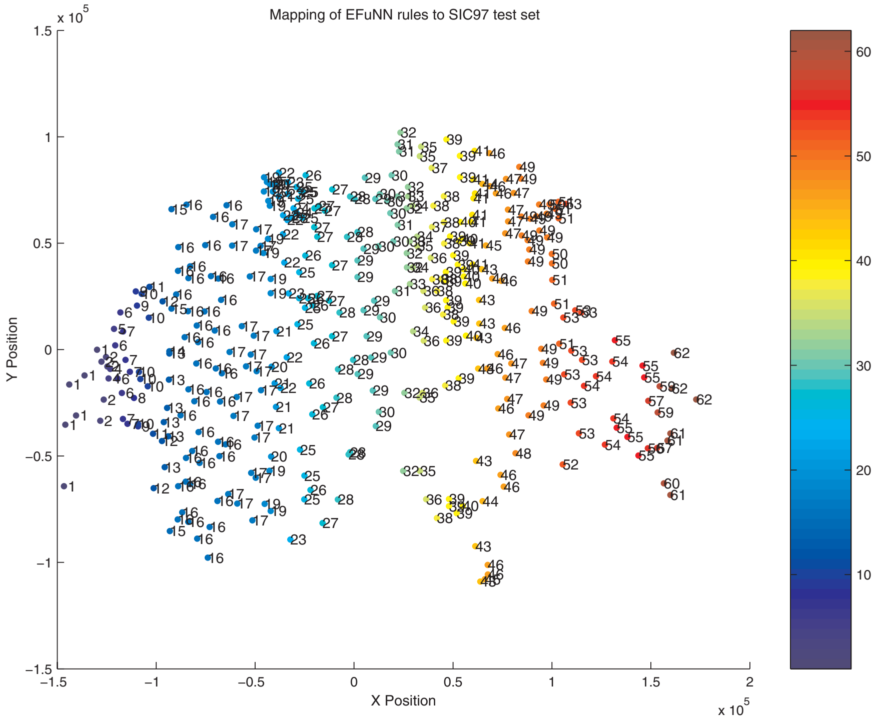


Figure 5: *Rule nodes mapped to testing data set*

4 CONCLUSIONS

In this paper we have proposed a new model, the EFuNN, as an alternative to the SVM for the purpose of local learning when applied to the SIC97 data set. In addition, we have shown that with the rule extraction ability of EFuNN enables us to better investigate how local learning impacts on such data sets by visualising a mapping of the EFuNN rule nodes on the data instances of the SIC97 training and testing data sets.

Future work will investigate how altering the number of rule nodes impacts on the performance of the EFuNN in order to create a more compact set of rules that better describe the dynamics of the data set. These rules can then be compared against the results other rule-extraction algorithms such as NeuroLinear (Setiono & Liu 1997).

References

Amari, S. & Kasabov, N. (1997). *Brain-like computing and intelligent information systems*. first edn. Springer Verlag.

Carpenter, G. & Grossberg, S. (1991). *Pattern Recognition By Self Organizing Neural Networks*. first edn. MIT Press: Cambridge, MA.

Chawla, S., Shekhar, S., Wu, W. & Ozesmi, U. (2001). “Modeling Spatial Dependencies for Mining Geospatial Data: An Introduction” In H. J. Miller & J. Han (eds), *Geographic Data Mining and Knowledge Discovery*. Taylor and Francis.

Cortes, C. & Vapnik, V. (1995). “Support-Vector Networks” *Machine Learning*. **20**(3): 273–297.

Dubois, G., Malzekski, J. & DeCort, M. (1998). “Spatial Interpolation Comparison” *Journal of Geographic Information and Decision Analysis*. **2**(2).

Gilardi, N. (2002). “Local Machine Learning Models for Spatial Data Analysis” *Geographical Information and Decision Analysis*. **4**(1): 11–28.

Kasabov, N. (1998a). “ECOS: A Framework For Evolving Connectionist Systems and the ECO Learning Paradigm” *Proceedings of ICONIP’98, Kitakyushu, Oct 1998*. Ohmsha, Ltd: Tokyo, Japan. pp. 1232–1236.

Kasabov, N. (1998b). “Evolving Fuzzy Neural Networks - Algorithms, Applications and Biological Motivation” *Proceedings of Iizuka’98, Iizuka, Japan*. World Scientific. pp. 271–274.

Kasabov, N. (1998c). “Fuzzy Neural Networks, Rule Extraction and Fuzzy Synergistic Reasoning Systems” *Research and Information Systems*. **8**: 45–59.

Kasabov, N. (1999a). *Evolving Connectionist And Fuzzy Connectionist System For On-Line Decision Making And Control*. Vol. Soft Computing in Engineering Design and Manufacturing Springer-Verlag.

Kasabov, N. (1999b). *Evolving Connectionist and Fuzzy-Connectionist Systems: Theory and Applications for Adaptive, On-line Intelligent Systems*. Vol. Neuro-fuzzy Tools and Techniques for Intelligent Systems, N Kasabov and R Kozma (eds) first edn. Springer-Verlag.

Kasabov, N. (2001). "On-line learning, reasoning, rule extraction and aggregation in locally optimized evolving fuzzy neural networks" *Neurocomputing*. **41**(1–4): 25–45.

Kasabov, N., Kim, J. S., Watts, M. & Gray, A. (1996). "FuNN/2- A Fuzzy Neural Network Architecture for Adaptive Learning and Knowledge Acquisition" *Information Sciences - Applications*. **101**(3-4): 155–175.

Kasabov, N. & Woodford, B. (1999). "Rule insertion and rule extraction from evolving fuzzy neural networks: algorithms and applications for building adaptive, intelligent expert systems" *Proceedings of the 1999 IEEE Fuzzy Systems Conference*. Vol. 3 The IEEE Kyunghee Printing Co. pp. 1406–1411.

Kohonen, T. (1990). "The Self-Organizing Map" *Proceedings of the IEEE*. **78**(9): 1464–1497.

Kohonen, T. (1997). *Self-Organizing Maps*. second edn. Springer-Verlag.

Ma, J. & Ahalt, S. (2001). "OSU SVM Classifier Matlab Toolbox (ver 2.00)".

Mitchell, M. T. (1997). *Machine Learning*. MacGraw-Hill.

Rumelhart, D. E., Hinton, G. E. & Williams, R. J. (1986). *Parallel Distributed Processing, Vols 1 and 2*. The MIT Press: Cambridge, MA.

Sammon, J. W. (1969). "A Nonlinear Mapping for Data Structure Analysis" *IEEE Transactions on Computers*. **18**: 401–409.

Setiono, R. & Liu, H. (1997). "NeuroLinear: from neural networks to oblique decision rules" *Neurocomputing*. **17**(1): 1–24.

Smola, A. J. & Schölkopf, B. (1998). "A Tutorial on Support Vector Regression" *Technical Report NC2-TR-1998-030*. NeuroCOLT 2.

Zadeh, L. (1965). "Fuzzy Sets" *Information and Control*. **8**: 338–353.