Monograph:
Downs, J., Harrison, R.F., Kennedy, R. Lee. et al. (1 more author) (1995) The Use of Fuzzy ARTMAP and Modified Fuzzy ARTMAP to Identify Low Risk Coronary Care Patients. Research Report. ACSE Research Report 556 . Department of Automatic Control and Systems Engineering

The Use of Fuzzy ARTMAP and Modified Fuzzy ARTMAP to Identify Low Risk Coronary Care Patients

Joseph Downs, Robert F Harrison
Department of Automatic Control and Systems Engineering
The University of Sheffield

R Lee Kennedy
Department of Medicine
The University of Edinburgh

Kent Woods
Department of Pharmacology
The University of Leicester

Research Report #556
12 January 1995

Abstract

The performance of fuzzy ARTMAP and modified fuzzy ARTMAP is compared using real-world data from a medical domain, the task being to predict the death or survival of patients admitted to a coronary care ward. Modified fuzzy ARTMAP is shown to perform consistently more accurately than fuzzy ARTMAP and is also much less prone to variations in performance with different orderings of training data. However, modified fuzzy ARTMAP does not show as large an improvement in performance as fuzzy ARTMAP when employed in the voting strategy. When unanimous voting decisions alone are considered, fuzzy ARTMAP is able to increase significantly accuracy in identifying survivors at the cost of decreased coverage of cases. This allows the identification of a subset of patients who have a low-risk of death from their condition and are thus potentially suitable for early discharge from hospital.

Correspondence Address

R.F. Harrison
Department of Automatic Control and Systems Engineering
The University of Sheffield
PO Box 600, Mappin Street
Sheffield, S1 4DU
United Kingdom

Telephone:+44 (0)114 2825139
Facsimile: +44 (0)114 2780409
E-mail: r.f.harrison@sheffield.ac.uk

1 Introduction

Since the fifties, there has been a progressive reduction in the recommended length of hospital stay for patients admitted to coronary care units. This change from six weeks bed rest to the current 7 to 10 days has occurred without significant effects upon patient mortality rates, and with obvious economic benefits. It is hoped to continue this trend yet further. However, early hospital discharge requires very accurate identification of those patients at minimal risk of death from their condition. Furthermore, such identification must occur soon after a patient is admitted. We have adopted a neural network approach to the task. To this end, a database of approximately 5500 patient records has been collected from the coronary care unit of Leicester Royal Infirmary over a six year period. Once records with incomplete data were removed, 4200 data items remained. Each record consisted of 43 items of clinical or electrocardiographic data considered to be useful for patient prognosis, together with the outcome for the patient's stay in hospital—death or survival. One of the items had an integer value, all others were binary-valued. (Appendix 1 provides a full listing of the input features.)

The generation of accurate predictions from this set of patient records is a very difficult learning task for a neural network. Specifically, two features of the data make the problem hard to solve. Firstly, the data is "noisy"—it is gathered from a real-world medical domain and has no simple indicators delineating category boundaries. Secondly, the distribution of categories is skewed—only 7.1% of all patients admitted die while on the ward. This poses particular difficulties for learning algorithms such as backpropagation (Rumelhart, Hinton and Williams, 1986) where weights are refined by a process which effectively averages together similar cases. In such circumstances rare events like patient deaths become completely submerged by the greater numbers of surviving patients who show similar data features.

2 Fuzzy ARTMAP and Modified Fuzzy ARTMAP

Fuzzy ARTMAP (Carpenter et al., 1992) was selected for the application since it is claimed to possess a number of capabilities that are particularly suited to this domain. First, fuzzy ARTMAP does not perform optimization of an objective function and is not therefore prone to the problem of local minima. Instead it forms a structuring of the data for itself (self-organisation) into prototypical category clusters. Also, fuzzy ARTMAP has few user-changeable parameters, which allows the model to be tuned to a particular problem without undue effort. (The single most important fuzzy ARTMAP parameter being that of *vigilance*, which controls the size of the category clusters formed). Additionally, a modified version of fuzzy ARTMAP (Lim and Harrison, In Press) has been demonstrated to show optimal data classification in the Bayesian sense.

Very importantly, the model is able to discriminate rare events from a "sea" of similar cases with different outcomes. This is because the family of adaptive resonance theory (ART) models to which fuzzy ARTMAP belongs all incorporate a feedback mechanism based on top-down matching of learned categories to input patterns (Carpenter and Grossberg, 1991). Learning in fuzzy ARTMAP can also occur with only one pass through the data set (single-epoch training). Furthermore, the model is capable of incorporating new data items without degradation of performance on previous data, or the necessity of retraining on such past data. This solution to the so-called *stability-plasticity dilemma* is claimed to be a feature unique among neural networks to the ART models (Carpenter and Grossberg, 1988). Collectively these features make fuzzy ARTMAP suitable for on-line learning, even in non-stationary environments.

Modified Fuzzy ARTMAP (Lim and Harrison, In Press) was developed from fuzzy ARTMAP for use with statistical data—the demonstration task being to separate two classes of Gaussian distributed random variables. The model was shown to provide superior performance to fuzzy ARTMAP on this problem, and can approach the Bayes optimal classification rates for the domain. To achieve this, modified Fuzzy ARTMAP records data on the usage of category cluster nodes, and resolves ties between nodes competing to encode an exemplar by selecting the node with the greatest usage.

3 Method

For the purposes of this application off-line learning was employed. The data were partitioned into a training set, comprising the first 3000 patient records, and a test set comprising the remaining 1200 records. Twenty different orderings of the training set were derived and served as input data to separate instances of fuzzy ARTMAP and modified fuzzy ARTMAP using single-epoch training. (The order of presentation of data items is known to have quite large effects on category formation in fuzzy ARTMAP, Carpenter et al., 1992.) The vigilance parameter was set low (0.3) to avoid excessive cluster formation—a notable problem for fuzzy ARTMAP (Marriott and Harrison, In Press). Other parameters were set to their "standard" values (see Kasuba, 1993); the learning rate being set to its maximum value of one (so-called fast learning) and the category choice parameter being set close to zero (0.000001).

The voting strategy (Carpenter et al., 1992) was also employed on the test data. This works as follows: a number of ARTMAP networks are trained on different orderings of the input data. During testing, each individual network makes its prediction for a test item in the normal way. The number of predictions made for each category is then totalled and the one with the highest score (or the most "votes") is the final predicted category outcome. The voting strategy can provide improved ARTMAP performance in comparison with the individual networks. In addition it also provides an indication of the confidence of a particular prediction, since the larger the voting majority, the more certain is the prediction.

A range of 3 to 13 odd numbered voters was used (odd numbers ensuring no tied decisions occurred), choosing those fuzzy ARTMAP instances from the pool of 20 that had achieved the highest individual accuracy scores. The same voting procedure was then repeated using the modified fuzzy ARTMAP networks.

4 Results

Initial performance on the test set proved disappointing. Accuracy for the individual fuzzy ARTMAP networks ranged between 73.2% and 87.5% with a mean of 81.1%. For modified fuzzy ARTMAP, accuracy ranged between 87.3% and 89.6% with a mean of 88.3%. (Full details are given in Appendix 2.) This compares with a default accuracy of 92.9% for the simple assumption that all patients will survive. The reason for this was that fuzzy ARTMAP in particular over-represents the rare cases of patient deaths in excess of their actual frequency within the data set. (This is probably because such cases were not tightly clustered together but widely spread throughout the feature space.) This effect was reduced, if not entirely overcome, with modified fuzzy ARTMAP. Thus fuzzy ARTMAP appears to suffer from the opposite problem to backpropagation—too much credence, rather than too little, is given to rare cases.

The general effect of the voting strategy was to increase accuracy for both fuzzy ARTMAP and modified fuzzy ARTMAP to around 89-91%, still slightly below baseline performance. However, the voting strategy with fuzzy ARTMAP did provide useful results for the important special case of high-confidence predictions of patient survival. (A high confidence prediction being one upon which all fuzzy ARTMAP voters agreed.) Such patients are the most suitable for early hospital discharge.

With 3 voters, a unanimous survival decision accounted for 911 of the data items and was proved wrong 44 times. This translates to 95.2% accuracy covering 75.9% of the 1200 test items. With extra voters, accuracy steadily improved at the cost of decreased coverage (see table 1 below), until at the 13 voter case an accuracy of 99.3% covering 34.0% of the data was achieved.

Table 1: Voting Strategy Performance for Unanimous Survival Decisions

	Fuzzy ARTMAP		Modified Fuzzy ARTMAP	
Number of Voters	Accuracy (%)	Coverage of Cases (%)	Accuracy (%)	Coverage of Cases (%)
3	95.2	75.9	94.0	87.0
5	95.6	64.5	94.2	81.3
7	97.8	53.0	94.8	75.3
9	98.2	45.3	95.1	71.8
11	98.1	40.3	94.9	69.1
13	99.3	34.0	95.1	66.3

5 Discussion

The results for the individual networks show that modified fuzzy ARTMAP consistently performs better than fuzzy ARTMAP with this real-world data. Moreover, the modified fuzzy ARTMAP networks show a much smaller variation in accuracy for different orderings of the training data. However, modified fuzzy ARTMAP does not gain as much benefit from the voting strategy as fuzzy ARTMAP. This is particularly marked when unanimous votes alone are considered (see table 1). With fuzzy ARTMAP, an increase in the number of voters tends to increase accuracy while reducing the number of cases. This effect is much less pronounced in modified fuzzy ARTMAP

Collectively these findings seem to indicate that modified fuzzy ARTMAP is not prone to the ordering effects of training data that occur with fuzzy ARTMAP. Thus modified fuzzy ARTMAP tends to form similar category clusters regardless of the order in which exemplars are presented. A single modified Fuzzy ARTMAP network therefore provides more reliable pattern classification than a Fuzzy ARTMAP network, and hence seems better suited to on-line learning tasks for example.

Appendix 1: Inputs to the Networks

1. Atrial Fibrillation
2. Supraventricular Tachycardia
3. Ventricular Tachycardia or Ventricular Fibrillation
4. Bundle Branch Block
5. ST Depression—Inferior
6. ST Depression—Anteroseptal
7. ST Depression—Anterolateral
8. ST Elevation—Inferior
9. ST Elevation—Anteroseptal
10.ST Elevation—Anterolateral
11.T Wave—Inferior
12.T Wave—Anteroseptal
13.T Wave—Anterolateral
14.Q Wave—Inferior
15.Q Wave—Anteroseptal
16.Q Wave—Anterolateral
17.Age<40
18.Age 40-50
19.Age 50-60
20.Age 60-70
21.Age>70
22.Time since Cardiac Event>24 Hours
23.Family History of Ischaemic Heart Disease
24.Smoker
25.Ex-Smoker
26.Glucose>7.9
27.Glucose>8.9
28.Glucose>10.9
29.Pulmonary Venous Engorgement
30.Pulmonary Oedema
31.Chest Pain
32.Short of Breath
33.Syncope
34.Nausea
35.Sweating
36.Palpitations
37.Cardiac Arrest while in Coronary Care Unit
38.Creatine Kinase>600
39.Creatine Kinase>1000
40.Creatine Kinase>3000
41.Sex
42.Number of Previous Cardiac Episodes
43.Previous Cardiac Episode(s)

Appendix 2: Individual Network Performance of Fuzzy ARTMAP and Modified Fuzzy ARTMAP

Table 2: Comparative Performance of ARTMAP and Modified Fuzzy ARTMAP

	Accuracy (%)	
Training Set Number	Fuzzy ARTMAP	Modified Fuzzy ARTMAP
1	73.2	89.4
2	86.2	87.7
3	85.1	88.0
4	81.0	89.3
5	78.5	87.3
6	74.8	89.3
7	79.0	89.6
8	83.1	88.6
9	79.8	88.8
10	87.5	88.5
11	83.2	87.7
12	84.2	88.2
13	82.3	89.3
14	78.3	87.8
15	78.4	88.3
16	84.3	87.7
17	82.5	87.8
18	80.0	87.8
19	79.6	88.5
20	81.8	87.4
Mean	81.1	88.3