Entropy Nets: From Decision Trees to Neural Networks

INTRODUCTION:

A multiple-layer artificial network (ANN) structure is capable of implementing arbitrary input-output mappings. Similarly, hierarchical classifiers, more commonly known as decision trees, possess the capabilites of generating arbitrarily complex decision boundaries in an n-dimensional space. Given a decision tree, it is possible to restructure it as a multilayered neural network. The objective of this project is to show how this mapping of decision trees into a multilayer neural network structure can be done. Finally, a two-step methodology for designing entropy networks is presented. The advantages of this methodology are that it specifies the number of neurons needed in each layer, along with the desired output. This leads to a faster progressive training procedure that allows each layer to be trained separately.

Method:

Classification using decision tree is performed by traversing the tree from the root node to one of the leaf nodes using the unknown pattern vector. The response elicited by the unknown pattern is the class or decision label attached to the leaf node that is reached by the unknown vector. It is obvious that all the conditions along any particular path from the root to the leaf node of the decision tree must be satisfied in order to reach that particular leaf node. . Thus, each path of a decision tree implements an AND operation on a set of half spaces. If two or more leaf nodes result in the same action or decision, then the corresponding paths are in an OR relationship.

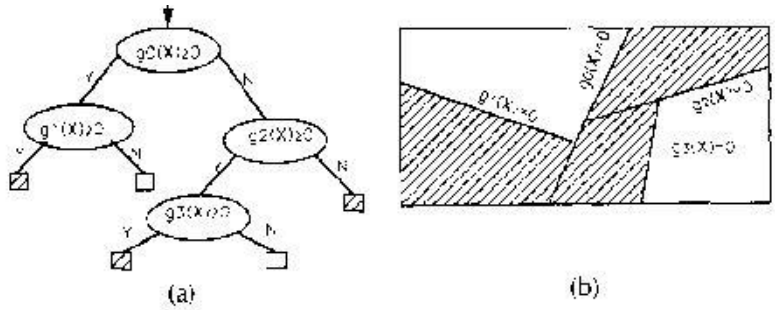

Fig. 2. (a) An example of a decision tree. Square boxes represent terminal nodes. (b) Hierarchical partitioning of the two-dimensional space induced by the decision tree of (a).

. Since a layered neural network for classification also implements ANDing Of hyper- planes followed by ORing in the output layer, it is obvious that a decision tree and a layered network are equivalent in terms of input-output mapping. Not only that a decision tree can be restructured as a layered network by following certain rules. . These rules can be informally stated as follows.

- The number of neurons in the first layer of the layered network equals the number of internal nodes of the decision tree. Each of these neurons implements one of the decision functions of internal nodes. This layer is the partitioning layer.
- All leaf nodes have a corresponding neuron in the second hidden layer where the ANDing is implemented. This layer is the ANDing layer.
- The number of neurons in the output layer equals the number of distinct classes or actions. This layer implements the ORing of those tree paths that lead to the same action.
- The connections between the neurons from the partitioning layer and the neurons from the ANDing layer implement the hierarchy of the tree.

An example of tree restructuring following the above rules is shown in Fig. 3 for the decision tree of Fig. 2. As this,

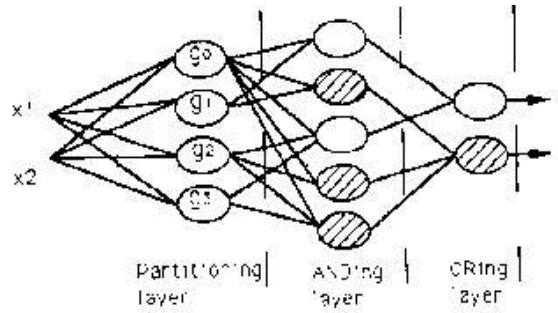

Fig. 3. Three-layered mapped network for the decision tree of Fig. 2(a).

Based on the discussion thus far, the following steps are suggested for designing entropy nets of pattern recognition tasks.

- Divide the available set of input-output mapping examples in two parts: tree design set and network training set. This should be done when a large number of input-output examples is available. Otherwise, the complete set of examples should be used for tree design and network training.
- Using AMIG or a similar recursive tree design procedure, develop a decision tree for the given problem. Map the tree into a three-layer neural net structure following the rules given earlier.
- Associate the sigmoid or some other soft nonlinearity with every neuron. Train the ANDing and ORing layers of the entropy network using the network training subset of the input-output mapping examples and the following procedure for determing the weight change.

CONCLUSION:

A new neural network design methodology has been presented in this project. This methodology has been developed by exploiting the similarities between the hierarchical classifiers of the traditional pattern recognition literature and the multiple-layer neural networks. It has been shown that the decision trees can be restructured as three-layer neural networks, called entropy networks