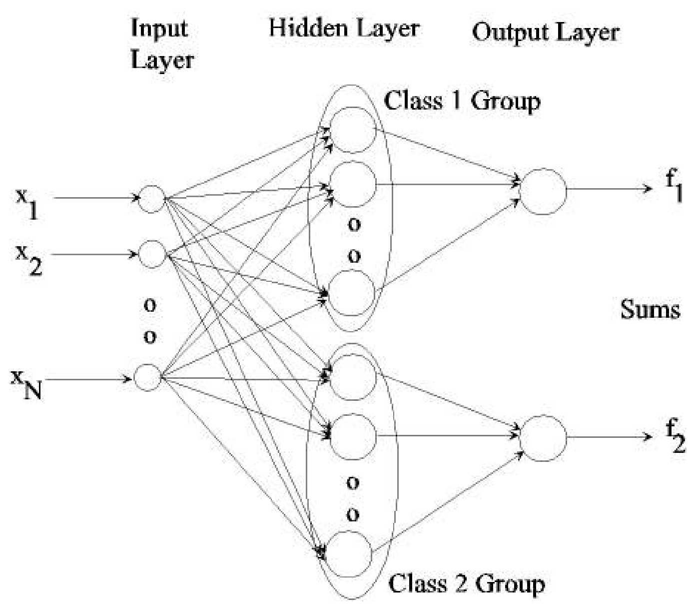
Fuzzy Neural Network Tutorial

Fuzzy Neural Networks

Our *fuzzy neural networks* (FNN's) are similar to the PNN's. Let there be K classes and let $\mathbf{x}$ be any feature vector from the population of interest to be recognized. The Class k exemplar feature vectors are denoted by $\mathbf{x}^{(q(k))}$ for q(k) = 1,...,Q(k). The summed functions here are not scaled and have a maximum value of unity.

$$f_1(\mathbf{x}) = (1/Q(1))\Sigma_{(q1=1,Q(1))} \exp\{-\|\mathbf{x} - \mathbf{x}^{(q(1))}\|^2/(2\sigma_1^2)\}$$
$$\vdots \qquad\qquad \vdots$$
$$f_K(\mathbf{x}) = (1/Q(K))\Sigma_{(q(K)=1,Q(K)} \exp\{-\|\mathbf{x} - \mathbf{x}^{(q(K))}\|^2/(2\sigma_K^2)\} \tag{1}$$



How the FNN Works

The summed functions in Equation (1) are averages of values between 0 and 1 and so are between 0 and 1. *Fuzzy logic* uses truth values between 0 and 1, so the output values $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ are the fuzzy truths that the input vector belongs to Class 1 and Class 2, respectively. We say the fuzzy truths are the values of *fuzzy set membership functions* whose functional values are the fuzzy truths of memberships each of the classes.

Thus the feature vector $\mathbf{x}$ belongs to the class with the highest fuzzy truth value (analogous to the highest value of a PNN). When there is a clear winner, then $\mathbf{x}$ belongs to a single class, but otherwise it may belong to more than one class with the given relative fuzzy truths. No training of weights is required. The feature vectors can be trimmed here and the spread parameters increased for efficiency.

An alternative way is to find the values for for all Gaussians, $f_1(\mathbf{x}),...,f_M(\mathbf{x})$, but not sum the ones for each class. Instead, we compute the values of the multiple Gaussians for each class and take the class maximum value as the output from the output node for that class.

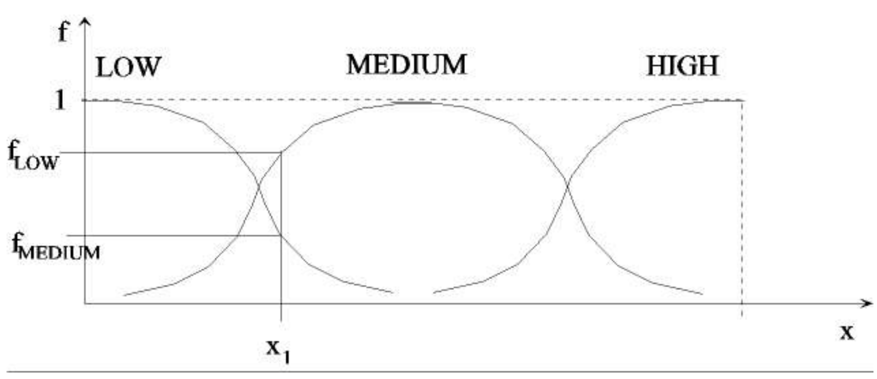
$$f_k(\mathbf{x}) = \max\ \{\exp[-\|\mathbf{x} - \mathbf{x}^{(q(k))}\|^2/(2\sigma_1^{\,2})]: 1 \le q(k) \le Q(k)\} \tag{2}$$

for each class k = 1,...,K. These K maximum values at the output nodes provide the fuzzy truth that the input feature vector belongs to each of the K classes. Then the maximum of all K of these values yields the winning class k* when it is significantly larger than any other values. Otherwise we can have multiple fuzzy set membership, that is, the input feature vector may belong to, say, Class 1 with a fuzzy value of 0.62 and it may also belong to Class 3 with fuzzy value of 0.57. If the data is good and not noisy, then one fuzzy set (class) will be a clear winner.

This alternative method of determining the fuzzy truths for each of the K classes assures that the fuzzy truth of membership in any class is between 0 and 1, whereas the first method above could possibly permit the fuzzy truth to exceed 1 (in which case we round it to 1).

Other Types of Fuzzy Neural Networks

There are many types of constructs for fuzzy NN's. The one described here uses fuzzy set membership function for each feature. For example, suppose the range of values for each of N features is the domain for 3 fuzzy set membership functions for 3 ranges of values respectively. Given the value x of a single feature from an input vector, we put it through the 3 fuzzy set membership functions, called LOW, MEDIUM and HIGH, for that feature. The figure below shows the 3 fuzzy set membership functions for the single feature x. Given a value x_1 we see that it yields the fuzzy truths of membership in the LOW fuzzy set with fuzzy truth $f_{LOW}(x_1)$ and a fuzzy truth of membership in MEDIUM with fuzzy truth $f_{MEDIUM}(x_1)$ (the f-axis is the fuzzy truth dimension).



Similarly, we do this for each of the other features. Then we have a set of fuzzy membership truths that are the values at the hidden nodes. These are compared to fuzzy truths of the features in the fuzzy sets for each class. There are many possibilities for doing this. One way is to let each feature vote for the class whose fuzzy set membership is closest with the majority vote the winner in determining the class. The advantage of this is that if there is noise on a minority of feature measurements it doesn't determine the winner. Other schemes are used and many more are possible.