ADAPTIVE NEURO FUZZY INFERENCE SYSTEM
FOR HIGHWAY ACCIDENTS ANALYSIS

Gianluca Dell’Acqua, Renato Lamberti e Francesco Abbondanti
Dept. of Transportation Engineering “Luigi Tocchetti”, University of Naples “Federico II”
E-mail: gdellacq@unina.it

ABSTRACT

In this paper we describe a method to model highway accidents using an adaptive neuro fuzzy inference system (ANFIS). Using a given input/output accidents data set we obtain a fuzzy inference system (FIS) whose membership function parameters are tuned using an optimization algorithm. This allows fuzzy system to learn from the data he is modelling. That is, we use ANFIS to train a FIS model to emulate the data presented to it by modifying the membership function parameters according to a chosen error criterion. Input variables include geometric and traffic functional parameters concerning the italian highway system. Accident rates are the output.

1 INTRODUCTION

The aim of this paper is to show how recent mathematical tools such as the fuzzy inference system and the neural network can be used in the study of road accidents, in order to obtain a model able to forecast an accident index, starting from the values of some input variables jointly regarded.

The distinctive feature of the fuzzy inference system is its underlying fuzzy logic, which provides a well suited approach for dealing with ill-defined and uncertain systems, with no need of conventional mathematical tools like differential equations. These latter are often unknown, and the effort to find them might be unrewarded when applying. In fact some quantity cannot be well expressed by means of numerical parameters, but linguistic (e.g., good, bad, little, big, much, few).

Furthermore, the verbose IF-THEN format that replaces the conventional mathematical equations, nearby the common human reasoning, allows to jointly relate several variables to another one.

The search of the non explicit mathematical relations between input variables and that of output, that is to say the “mapping” of the input-output space, is performed using a data set; the co-ordinates of each data point are fixed values of the input and output variables. That’s why for giving the inference fuzzy system the adjective adaptive (that is tuned on data). The learning process consists of two steps: in the second step takes place the “training” of a neural network. Owing to this, the further adjective neuro.
2 METHODOLOGY

Fuzzy logic is a mathematical-linguistic approach trying to describe analogic quantity fluctuation more efficiently than conventional logic.

Fuzzy logic extends traditional set theory relaxing the ties of non contradiction and excluded third principles: an element can belong to a set and not belong at the same time.

A fuzzy set is represented by a pair:

\[ A = \{ x, \mu_A(x) \} \]

in which \( \mu_A(x) \) is a membership function associating each element \( x \) to a real number \( \in [0, 1] \).

Intersection and union logic operators used in the classic set theory, have a fuzzy extension in the corresponding operators and or:

\[ \mu_{A \text{ and } B} = \min[\mu_A(x); \mu_B(x)] \]

\[ \mu_{A \text{ or } B} = \max[\mu_A(x); \mu_B(x)]. \]

2.1 Fuzzy Inference System

In general terms it is a fuzzy logic based process which provides an output starting from some inputs. The typical scheme consists of the following steps.

2.1.1 Fuzzyfication

In this step deterministic values of the input variables are turned in membership degree to fuzzy sets. These sets are labelled with commonly used linguistic values.

Conversion of deterministic values into linguistic values is carried out by membership functions; these latter fix a relation between each deterministic value and a degree of matching with one or more linguistic values chosen for a quantity, in a scale ranging from 0 to 1, as shown in Figure 1:

\[ \begin{align*}
A & \quad B \\
\mu_A(x°) & \quad \mu_B(x°)
\end{align*} \]

Figure 1. Membership functions

2.1.2 Applying Rules

Each rule establishes a relation between the linguistic values through an IF-THEN statement:
IF $x_1$ is $A_{j1}$ AND ... AND $x_i$ is $A_{ji}$ AND ... AND $x_n$ is $A_{jn}$ THEN $y$ is $B_j$

where $x_{i=1...n}$ are the input variables, $y$ is the output variable, $A_{j,i=1...n}$ and $B_j$ are linguistic values labelling fuzzy sets. The degree with which the output variable $y$ matches the corresponding fuzzy set (his linguistic value) $B_j$, depends on the degrees of matching of the input variables $x_{i=1...n}$ to theirs fuzzy sets $A_{j,i=1...n}$ and on the logic format (AND, OR) of the antecedent part of the rule. So, it is immediate calculating the degree of matching in each rule as shown in Figure 2:

![Figure 2. Applying rules](image)

2.1.3 Logic Sum and Defuzzyfication

Each rule gives as a result a fuzzy set, with a membership function cut in the higher zone. By all the rules is given a set of fuzzy sets with differently cut membership functions, whose deterministic values all have a share in the inferential result. A single value is needed in order to have an useful result. The most commonly employed method fixes this value to the barycentric abscissa of the area resulting from the joining the cut functions (Figure 3).

![Figure 3. Logic sum and defuzzyfication](image)
2.2 Tuning Membership Functions By Means Of A Data Set

A data set which is regarded to be well descriptive of the phenomenon, can be used to fix the membership function parameters.

If membership functions are gaussian-type (as it often is), the chosen logic operator is the multiplication, and the defuzzification method is the weighted average of maximums, then the output can be calculated as:

\[ y = f(x) = \frac{\sum_j y_j^c \prod_i \mu_{A_{ji}}(x_i)}{\sum_j \prod_i \mu_{A_{ji}}(x_i)} \]

where

\[ \mu_{A_{ji}} = e^{-\left(\frac{(x_i - x_{ji}^c)^2}{\sigma_{ji}^2}\right)} \]

is the degree of membership of the i-th input quantity to the fuzzy set A_{ji}.

Once fixed the values of the i input quantity \( x_i = 1 \ldots n \) the output value \( y \) depends on the values of the centres of the output membership functions \( y_j^c \) and of the centres \( x_{ji}^c \) and parameters \( \sigma_{ji} \) of the input membership functions; if input-output values (\( x_k, y_k \), \( k = (1, K) \)) of a sample data set are available, the values of \( y_j^c, x_{ji}^c \) and \( \sigma_{ji} \) can be forced to be such that the difference between the outputs provided by the system \( y = f(x_k) \) and those provided by the data set \( y_k \), is the minimum.

This can be obtained applying the well known gradient descent optimization method to the following objective function:

\[ c_k = (f(x_k) - y_k)^2 / 2. \]

2.3 Fuzzy Clustering

Fuzzy Clustering is a classification method for sample data.

Consider \( X = (x_1, x_2, \ldots x_n) \) a collection of \( n \) data points in an \( M \)-dimensional space and bounded by a hypercube (that is normalized in each dimension). The aim is to share the points into \( c (2 \leq c \leq n) \) clusters so as the points belonging to a same cluster are much more similar than the points belonging to different clusters. As similarity measure it’s obvious choosing the distance between pairs of points in the \( M \)-dimensional space.

The used method, known as fuzzy c-means, is due to Bezdek (1981) and is a generalization of the hard c-means method for the crisp classification.

2.4 Building A Fuzzy Inference System By Means Of A Data Set

As well as for the tuning of the membership functions parameters, a sample data set can be used to generate the rule base in terms of number and syntax of each rule, that is to obtain the structure of the fuzzy system model.

This is made through a clustering of the data. In fact if the j-th rule has the form:

IF \( x_1 \) is \( A_{j1} \) AND ... AND \( x_i \) is \( A_{ji} \) AND ... AND \( x_n \) is \( A_{jn} \) THEN \( y \) is \( B_j \)
it is easy to think at it like a cluster: a data point would be more or less near the cluster center as well as the values of its variables would be more or less matching the linguistic values labelling the fuzzy sets $A_{j,i=1...n}$ and $B_j$.

The cluster centers coordinates give the initial estimate for the membership function parameters, then to be tuned.

Different methods differ in clustering way and in tuning (optimization) way.

In ANFIS (Jang (1993)) the fuzzy model can be also represented with a neural network structure.

3 APPLICATION: THE ITALIAN HIGHWAY SYSTEM

We analysed 53 Italian highway-sections with toll-gate for which quarterly indexes about traffic, covered distances and accidents were available.

Then we chose 7 input variables. 5 of them represent geometrical factors: width of the road, width of the emergency-lane, average distance between tunnel, average distance between bridges, tortuosity index. The remaining 2 represent traffic factors: number of daily entered vehicles, heavy vehicles ratio.

As output variables we chose two accident indexes: the accidents rate (number of accidents related to the total distance covered by all entered vehicles) and the accidents number (obviously related to the length of each section).

Considering one of these output variables at a time, we built two data sets in order to obtain two fuzzy models able to forecast for the respective accident indexes. To accomplish this task we used the Fuzzy Logic Toolbox belonging to MATLAB’s environment (The Math Works Inc.).

Each data set is at first used to establish the rule base through a clustering algorithm; then the same data set is used to obtain the membership functions parameters through an optimization algorithm.

Once obtained the FIS (fuzzy inference system), it proves its usefulness when using the rules: their “black box” function allows to instantly obtain the output value simply giving the input values (Figure 4).

During the phases of elaboration it was possible to give a judgement about the degree of reliability of each obtained model in terms of a correct simulation of the phenomenon; well, the degree of reliability resulted to be much more high for the model forecasting the accidents number respect of the one forecasting the accidents rate.

Being the 7 input variables the same in the two models, clearly appears the importance of the kind of these variables, in addition to their number, to obtain a good modeling of the phenomenon.
4 CONCLUSIONS

We think that ANFIS can be another efficient mathematical tool, besides those of the traditional binary logic, to deal with the study of road risk. For the sake of reducing this risk, such an accidents-modeling can help in singling out the corrective actions, whether in terms of improving the road characteristics or in terms of reorganizing the road network supply.

REFERENCES
