Chapter 6

EXTRACTING INTERPRETABLE FUZZY KNOWLEDGE FROM DATA

A Neuro-Fuzzy model to extract human understandable symbolic knowledge from numerical examples

CORRADO MENCAR

Università degli Studi di Bari – Dipartimento di Informatica

6.1 INTRODUCTION

A Fuzzy Knowledge Base (FKB) is a knowledge base composed of fuzzy rules. It is able to express – in a human readable form – vague properties about an environment and allows making predictions about new environmental stimuli.

Learning fuzzy knowledge is the process to discover a set of fuzzy rules from available data that are able to explain relationships among data features and predict them for future data. Neuro-Fuzzy models are connectionist learning machines capable of acquiring knowledge from data and representing it in form of fuzzy rules. The main issue that arises with such models is interpretability of the discovered FKB, which is heavily compromised if no special attention is paid during model training.

In order to extract knowledge to be judged “interpretable”, a set of properties on FKBs must be formulated. A lot of work has been done in this sense (Pedrycz & Gomide, 1998; de Oliveira, 1999; Chow, Altug & Trussel, 1999), even if to date there is no well-established definition for interpretability of a FKB.

Several approaches have been proposed to obtain interpretable knowledge by neuro-fuzzy learning, like those in (de Oliveira, 1999; Chow et al., 1999; Nauck, Nauck & Kruse, 1996). Many of such approaches,
however, require great computational efforts in order to achieve interpretable FKBs.

Here, we get insight in an approach to automatically extract fuzzy rules by learning from data, with the main objective to obtain interpretable knowledge (Castellano, Fanelli & Mencar, 2002a). The approach is based on a neuro-fuzzy model designed so that its learning algorithm works in a parameter space with a reduced dimensionality. The dimensionality of the new parameter space is necessary and sufficient to generate human-understandable fuzzy rules, in the sense formally defined by a set of properties. Once the new parameter space is defined, the learning algorithm performs simple gradient descent with no additional constraints in the parameter modifications. The resulting model has the appreciable quality of avoiding extra-computations necessary for other models to achieve comprehensible FKBs. Moreover, the FKB of the model is always understandable, even during learning process.

6.2 THE FUZZY INFERENCE SYSTEM

Reasoning with fuzzy rules needs the formal definition of a Fuzzy Knowledge Base and a Fuzzy Inference System (FIS), which is a formal apparatus that derives fuzzy properties from the facts, i.e. the system inputs, and the fuzzy rules contained in the FKB.

6.2.1 The Fuzzy Knowledge Base

The rule base schema, defined in the following, is equivalent to the schema used in a 0-order TSK FIS (Jang, 1993):

\[
\text{Rule } r: \text{ IF } x_1 \text{ IS } A_1^{g(r,1)} \ldots \text{ AND } x_n \text{ IS } A_n^{g(r,n)} \text{ THEN } y_1 = v_{r,1}, \ldots, y_m = v_{r,m}, \quad r = 1, 2, \ldots, R
\]  

where \( n \) is the number of inputs and \( R \) is the total number of rules. The symbols \( A_i^{g(r,i)} \) denote input fuzzy sets with membership function \( \mu_{A_i^{g(r,i)}} \).

The function \( g : \{1, 2, \ldots, R\} \times \{1, 2, \ldots, n\} \to \mathbb{N} \) is used to share the same fuzzy sets in different rules. For a given rule \( r \) and an input \( i \), the index \( g(r,i) \) represents the fuzzy set of the \( i \)-th input variable used in the \( r \)-th rule.

Fuzzy sets \( A_i^{g(r,i)} \) are represented by Gaussian membership functions: