Chapter XVIII
Fuzzy Miner:
Extracting Fuzzy Rules from Numerical Patterns

Nikos Pelekis
Greece & UMIST Manchester, UK

Babis Theodoulidis
UMIST Manchester, UK

Ioannis Kopanakis
UMIST Manchester, UK

Yannis Theodoridis
University of Piraeus, Greece

INTRODUCTION

Recently, our capabilities of both generating and collecting data have increased rapidly. Consequently, data mining has become a research area with increasing importance. Data mining, also referred to as knowledge discovery in databases (Chen et al., 1996), is the search of relationships and global patterns that exist "hidden" among vast amounts of data. There are various problems that someone has to deal with when extracting knowledge from data, including characterization, comparison, association, classification, prediction, and clustering (Han & Kamber, 2001). This chapter elaborates on the problem of *classification*. Broadly speaking, pattern classification (or recognition) is the science that is concerned with the description or classification of measurements. More technically, pattern classification is the process that finds the common properties among a set of objects in a database and classifies them into different classes according to a classification model.

Classical models usually try to avoid *vague, imprecise*, or *uncertain* information, because it is considered as having a negative influence in

the inference process. This chapter accepts the challenge of dealing with such kind of information by introducing a fuzzy system, which deliberately makes use of it. The main idea of fuzzy systems is to extend the classical two-valued modeling of concepts and attributes like *tall, fast*, or *old* in a sense of gradual truth. This means that a person is not just viewed as *tall* or *not tall,* but as tall to a certain degree between 0 and 1. This usually leads to simpler models, which are handled more easily and are more familiar to the human way of thinking.

After providing a brief comparative overview of pattern classification approaches (Section 2) and a short specification of the pattern classification domain in fuzzy systems (Section 3), the chapter follows the above paradigm and describes an effective fuzzy system for the classification of numerical data (Section 4). The initial idea comes from the fact that fuzzy systems are universal approximators (Kosko, 1992; Wang, 1992) of any real continuous function. Such an approximation method (Nozzaki et al., 1997) coming from the domain of fuzzy control systems is appropriately adjusted, extended, and implemented in order to produce a powerful working solution in the domain of pattern classification. An "adaptive" process is also introduced, developed, and incorporated into the previous mechanism for automatically deriving highly accurate linguistic if-then rules. The description of the methodology is combined with the illustration of the design issues of the tool *Fuzzy Miner*. The current work is evaluated (Section 5) by extensive simulation tests and by providing a comparison framework with another tool of the domain that employs a neuro-fuzzy approach, NEFCLASS (Nauck & Kruse, 1995). Finally, the chapter concludes (Section 6) by identifying promising directions for future work pointed to by this research effort.

COMPARATIVE OVERVIEW OF PATTERN CLASSIFICATION APPROACHES

Already, when the field was still in its very infancy, it was realized that the statistics and probability theory (Berger, 1985) had much to offer to pattern classification (Schalkoff, 1992). The question of whether or not a given pattern "belongs" to some pattern class may naturally be treated as a special case of the statistical decision theory problem. Effective, though, as it is, the statistical approach has built-in limitations. For instance, the theory of testing statistical hypotheses entails that a clear-cut yes or no answer should always decide upon the membership of a pattern in a given class. Clearly, not all of the real life patterns admit of such coarse decisions. Sometimes information in a pattern is not simply in the presence or the absence of a set of features, but rather the interconnection of features contains important structural information. Indeed, this relational information is difficult or impossible to be quantified by a feature vector form. This is the underlying basis of structural pattern classification. Structurally based systems assume that pattern structure is quantifiable. As such, complex patterns can be decomposed recursively in simpler subpatterns in almost the same way that a sentence can be decomposed in words. The analogy directed researchers toward the theory of formal languages. The process that results in an answer to a classification question is called syntax analysis or parsing.

Fuzzy Logic and Fuzzy Systems for Pattern Classification

Fuzzy logic (Zimmermann, 1996) is a superset of conventional (Boolean) logic that has been extended to handle the concept of partial truth