Chapter XX
Fuzzy Classification in Shipwreck Scatter Analysis

Yauheni Veryha
ABB Corporate Research Center, Germany

Jean-Yves Blot
Portugal Institute of Archaeology, Portugal

Joao Coelho
Portugal Institute of Archaeology, Portugal

ABSTRACT
There are many well-known applications of fuzzy sets theory in various fields of science and technology. However, we think that the area of maritime archaeology did not attract enough attention from researchers of fuzzy sets theory in the last decades. In this chapter, we present examples of problems arising in shipwreck scatter analysis where fuzzy classification may be very useful. Using a real-world example of fragments of ceramics from an ancient shipwreck, we present an exemplary application of the fuzzy classification framework with SQL querying for data mining in archaeological information systems. Our framework can be used as a data mining tool. It can be relatively easily integrated with conventional relational databases, which are widely used in existing archaeological information systems. The main benefits of using our fuzzy classification approach include flexible and precise data analysis with user-friendly information presentation at the report generation phase.

INTRODUCTION
At the stage of the typical data classification in information systems, there appear some types of uncertainty, for instance, when the boundaries of a class of objects are not sharply defined (Borgodna, Leporati, Lucarella, & Pasi, 2000; Kacprzyk & Zadrozny, 2000). In this case, the most...
common, useful, and widely accepted approach is the introduction of data fuzzification (Bellma & Vojdani, 2000; Schindler, 1998). Fuzzy sets provide mathematical meanings to natural-language statements and become an effective solution for dealing with uncertainty. There is no simple procedure for combining various attributes that define a particular data set into one general performance measure because attributes may be measured with different scales, the relative significance of different criteria differs, and for some criteria the objective is maximization while for others it is minimization or another specific target. The approach of fuzzy sets theory with its membership functions is widely used to form a realistic aggregated description of the data set. In fuzzy sets theory, various attributes with separate scales and optimization objectives can be combined into a joint response measure: the aggregated value of membership (Zimmermann, 1992; Zimmermann and Zysno, 1980).

The precise analysis of shipwreck data in maritime archaeology leads us to distinguish between ship structure and equipment, ship cargo, and personal items present on board at the time of the accident. However, the data appear strongly blurred in shallow underwater sites marked by long-term sea dynamics (Blot, 1998; Muckelroy, 1978). In some cases, the only remains of the original shipwreck are composed of fragments of durable materials associated with the cargo, ceramics containers (amphorae for instance), or coins. In such an extreme environment where no remains of the ship itself have been found yet, and neither should it be expected due to the sea environment made of protruding rocks and shallow sandy patches, no clear pattern emerges from the immediate reading of the spatial distribution of the fragments. In the meantime, some physical parameters attached to the fragments themselves may be explored. The matter has been previously discussed within a purely quantitative approach related to coins in a shallow oceanic environment (Vargiolu, Zahouani, & Blot, 2005).

The case examined here deals with a strictly qualitative approach applied this time to broken ceramics collected from a shipwreck in Peniche, Portugal, that happened some two millennia ago, of a small vessel coming from southwest Spain (Roman province of Baetica) with a cargo of wine amphorae (Haltern 70) and some fine ceramics (Samian ware) from Italy. Three of the parameters isolated in the original examination of the ceramics materials (abrasion, cleavage, and neat fracture) have been selected and implemented using a purely qualitative scaling incorporating linguistic terms (e.g., visible, invisible, slightly visible, etc.) and experts’ opinions within a specially defined scale of points, which are later assigned to a particular shipwreck fragment. Those parameters find an echo in sedimentary petrology (Dobkins & Folk, 1970) or in hydrology with prehistoric tools. Shackley, long ago, illustrated with flint implements “how measurements of abrasion can indicate different environments, i.e. edge rounding by sand and smaller particles and percussion craters formed by impacts with larger pebbles” (Brown, 1997; Shackley, 1974).

Unlike hydrology where the search is directed toward the measurements of transport and fluxes, archaeology tends to look for clues related to “nonmovement” and physical parameters observable on the artifacts that may be relevant to the spatial analysis of a site. In the case commented on in this chapter, our interest was directed toward ceramic items brought from the underwater site that might be associated with the initial moments of the nautical accident that occurred some two millennia ago. The experiment was thus directed toward testing the three descriptive parameters referred to above within a group of several hundred ceramic sherds observed and described by three experts or judges acting independently. Among the several biases subjacent to the experiment was the learning curve related to the personal adjustment of each expert in using unconventional, previously untested, descriptive parameters.

Even when using classical statistical techniques, like principal component analysis or factorial analysis, which are very robust in mathematical terms, they often do not allow easily integrating