The Second World Landslide Forum Abstracts
WLF2 - 2011– 0208
Rome, 2011

Modeling landslides' susceptibility by fuzzy emerging patterns mining and evaluation

Anna Rampini (1) , Gloria Bordogna (2) , Monica Pepe (1) , Massimo Antoninetti (1) , Alessandro Mondini (3) , and Paola Reichebach (3)

(1) CNR, Irea, Milano, Italy; (2) CNR, Idpa, Milano, Italy; (3) CNR, Irpi, Perugia, Italy;

In this contribution an approach to model landslides susceptibility of a territory is proposed, based on a novel supervised learning technique that, on a set of classified data used as training data, identifies fuzzy emerging patterns. Many approaches have been proposed and applied in the literature, mainly based on statistical classification, physical landslides models, fuzzy logic. Fuzzy emerging patterns are a generalization of emerging patterns which are a combination of attribute values that occurs mostly in a class, which barely appear in the remaining classes; so the presence of a pattern in an item (an element or instance, of the database) gives some evidence about the class the object should belong to. Many authors propose algorithms to extract and use emerging patterns for supervised classification. Mainly two families of emerging pattern–based classifiers have been proposed: the first makes a lazy discovery, since in the classification stage, the classifier searches for patterns contained in the object ; in contrast, the second family, searches all emerging patterns in the training stage, and later uses them to classify. We use this second approach to identify fuzzy emerging patterns on training data. In our application two sets of fuzzy pattern are identified, a positive and a negative set. The positive set contains the fuzzy patterns that characterize the territory with respect to its disposition to landslide occurrences. Conversely, the negative set comprises the fuzzy emerging patterns that characterize the disposition to stability of the same territory. We mine these two sets of fuzzy emerging patterns separately from two subsets of the training data; the training data contains items or instances which correspond to units, i.e., homogeneous area of a given territory. The units are described by some thematic layers of information that constitute the attributes, and they assume a value for each attribute. Moreover each unit is classified into two classes (susceptible or not to landslides) depending on the percentage of its area that was actually affected by a landslide in the past. We divide the training data set into two subsets corresponding with the two distinct classifications. Once these patterns are identified we can use them to classify the territory. The classification associates with each unit two degrees, a degree of favour and a degree of disfavour to the occurrence of landslide. Based on these degrees we generate a susceptibility map to landslides and a certainty degree of this classification, that reflects the hesitation margin associated with the decision process. The originality of the proposal is twofold: 1) the identification of fuzzy emerging patterns allows building an explicit and comprehensible modelling of the factors that favour and disfavour landslide events, that can aid the expert geologist to integrate their knowledge on the phenomenon of landslides. 2) each basin is classified with a degree of susceptibility and an associated hesitation degrees that expresses how certain and indubitable is the obtained classification. For example two distinct basins can be classified with the same high degree of susceptibility s1=s2=0.8 but different hesitation margin, e1=0.9 and e2=0.1. This means that while we can be almost sure that basin 1 has a high susceptibility to landslide, for basin 2 we arrived to the same susceptibility but we doubt on its truthfulness, since the available mined knowledge, expressed by the fuzzy emerging patterns, is contradictory. These second situation must be tell apart by revising the available knowledge with new training data and new analysis.