Chapter 86
Managing Uncertainty in Geospatial Predictive Models

Iftikhar U. Sikder
Cleveland State University, USA

ABSTRACT

Geospatial predictive models often require mapping of predefined concepts or categories with various conditioning factors in a given space. This chapter discusses various aspects of uncertainty in predictive modeling by characterizing different typologies of classification uncertainty. It argues that understanding uncertainty semantics is a perquisite for efficient handling and management of predictive models.

1. SPATIAL PREDICTION AND CLASSIFICATION

Geospatial predictive models entail an array of analytical techniques of data mining, classical statistical and geostatistical models that attempt to predict spatial states and behavior of objects from a fine set of observations. The process of prediction presupposes a set of spatial concepts and categories to which objects are to be mapped. For example, spatial processes, such as classification of land cover from satellite image, modeling forest fire, propagation of epidemics, and prediction of urban sprawl require a unifying and common reference of "space" or location where the multiple features of spatial attributes are to be mapped to predefined class labels. The prediction of spatial features can be conceived as a process of driving classification schemes in relation to certain spatial properties such as neighborhood, proximity, dependency, as well as similarity of non-spatial attributes (Han & Kamber, 2006; Shekhar & Chawla, 2003). In data mining, a classification function is often defined as a mapping function: $f : A \rightarrow \mathrm{C}$, where A is the domain of function, f represents attribute space and C is the set of class categories.

DOI: 10.4018/978-1-4666-2038-4.ch086

2. UNCERTAINTY IN SPATIAL CLASSIFICATION

Uncertainty may emerge from ontological constraints in classification i.e., from the lack of specification of what kind of spatial objects exist, as well as from epistemic limitations which concern whether such objects are knowable to subjective schemes, and if so, to what extent they can be represented in the subjective framework, given the limited empirical evidences. Epistemic uncertainty in spatial classification emerges due to inadequate representation of spatial knowledge which is often incomplete, imprecise, fragmentary, and ambiguous. The attributes of spatial objects or evidences suggesting various conceptual or thematic classes may often suggest conflicting categories. Moreover, classification labels are dependent on the resolution of observation and the extent of granularity. For example, the observation of coarser granularity offers less detail while the clumping of information into pixels in remotely sensed images may prevent sub-pixel entities being distinguished (Fisher, 1997). The classification of land cover from satellite image depends not only on a specific spatial resolution, radiometric resolution and the corresponding spectral signatures limit predictive accuracy. Therefore, spatial characteristics of a given observation are indiscernible with respect to attributes associated with it. For example, the number of vegetation types that can be identified from an NDVI (Normalized Difference Vegetation Index) image significantly increases when a very high radiometric resolution is used. Moreover, in a specific case, a multispectral image may provide more accuracy than a hyperspectral image, but such accuracy is of little value if it is achieved at the cost of less specificity or higher imprecision.

3. TYPOLOGIES OF CLASSIFICATION UNCERTAINTY

While there is increasing awareness of uncertainty, and its aspects and dimensions in predictive as well as classificatory schemes, little agreement exists among experts on how to characterize them. Many typologies of uncertainty have been suggested from risk analysis perspective, which often overlaps and builds on each other (Ferson & R.Ginzburg,1996; Linkov & Burmistrov2003; Regan et al., 2002). These typologies make distinctions between variability and lack of knowledge at the parameter and model level. However, from the geographic information perspective, the ontological specification of imperfection of geographic data provides some key vocabularies and taxonomies to deal with spatial uncertainties (Duckham et al., 2001; Worboys & Clementini, 2001). Such ontology distinguishes between inaccuracy (i.e., errors or commission or omission) and imprecision, which arises from limitations on the granularity of the schema or levels of detail obtainable for an observation under which the observation is made (Worboys, 1998). The concept "vagueness" refers to indeterminate boundary-line cases or "inexact concepts".

Classification of geographic objects with indeterminate boundaries offers many challenges (Burrough & Frank, 1996) which emerge from the boundary of many real entities representing natural, social, or cultural phenomena (for example, forests, mountains, areas ethnic distribution etc.). Since many common geographical concepts are vague (Fisher, 2000), the explicit specification of vagueness is essential to characterize the classification performance. As a special type of vagueness, nonspecifity originates due to our inability to discern the true alternatives among several alternatives in a given context. It implies cardinality of undiscerned alternatives (Klir & Yuan, 1995). The larger the set of alternatives, the higher is the nonspecifity. For example, in a remotely sensed image, a pixel with class type "forest" and the mean annual temperature > 30^{C} has less nonspecifity than the pixel labeled only with "forest" type. This is because in the latter case a pixel can have a large number of possible variations of "forest" type.

Broadly, three major categories of uncertainty can be identified in dealing with predictive and classificatory problems: ontological uncertainty,

Related Content