Generative and Discriminative Facets of Granular Data

Witold Pedrycz

Abstract

Concepts constitute a concise manifestation of key features of data. As being built at the higher level of abstraction than the data themselves, they capture their essence and usually emerge in the form of information granules. Data descriptors – information granules characterize data in a succinct way and are further used as functional building blocks (modules) of classifiers or predictors.

Generative and discriminative aspects of information granules and the notion of information granularity deserve detailed investigations. Generative aspects of information granule are concerned about a descriptive representation of data by a collection of information granules. Discriminative aspects of granules deal with their abilities to realize mappings of classification (discrete) or predictive (continuous) character.

In this talk, a systematic process of design of information granules completed in the presence of numeric data is elaborated on. Two design phases are delineated. First, a collection of numeric representatives (prototypes) is formed with the aid of clustering algorithms. In the sequel, such representatives are augmented (elevated) to information granules with the help of the principle of justifiable granularity. Several variants of this principle are discussed to reflect the essence of the underlying problem, cope with available domain knowledge, and address key requirements of the problem. Linkages with AI and Granular Computing are identified and a symbolic aspect of information granules is identified.

The discriminatory nature of information granules is illustrated in the general setting of granular classifiers, aggregation schemes is carried out in collaborative and distributive environment, and with regard to the problem of transfer learning.

Furthermore we elaborate on the general paradigm of granular modeling dwelling upon a collection of information granules and discuss a phenomenon of a successive elevation of type of information granularity emerging through refinements of numeric and granular models.