Soft Computing, f-Granulation and Pattern Recognition

Sankar K. Pal

Indian Statistical Institute, Kolkata 700108, India

sankar@isical.ac.in

<http://www.isical.ac.in/~sankar/>

Abstract. Different components of soft computing (e.g., fuzzy logic, artificial neural networks, rough sets and genetic algorithms) and machine intelligence, and their relevance to pattern recognition and data mining are explained. Characteristic features of these tools are described conceptually. Various ways of integrating these tools for application specific merits are described. Tasks like case (prototype) generation, rule generation, knowledge encoding, classification and clustering are considered in general. Merits of some of these integrations in terms of performance, computation time, network size, uncertainty handling etc. are explained; thereby making them suitable for data mining and knowledge discovery.

Granular computing through rough sets and role of fuzzy granulation (f-granulation) is given emphasis. Different applications of rough granules and certain challenging issues in their implementations are stated. The significance of rough-fuzzy computing, as a stronger paradigm for uncertainty handling, and the role of granules used therein are explained with examples. These include tasks such as class-dependent rough-fuzzy granulation for classification, rough-fuzzy clustering, and defining generalized rough entropy for image ambiguity measures and analysis. Image ambiguity measures take into account the fuzziness in boundary regions, as well as the rough resemblance among nearby gray levels and nearby pixels.

Significance of rough granules and merits of some of the algorithms are described on various real life problems including multi-spectral image segmentation, determining bio-bases (c-medoids) in encoding protein sequence for analysis, and categorizing of web document pages (using vector space model) and web services (using tensor space model). The talk concludes with stating the possible future uses of the methodologies, relation with computational theory of perception (CTP), and the challenges in mining.