Future directions in unsupervised soft clustering

Pawan Lingras

Saint Mary's University, Halifax, Canada

Unsupervised learning

- Large datasets
- Difficult to get classifications for billions of objects
- Unsupervised grouping can be useful to simplify user surveys
- Unsupervised clustering can be combined with supervised learning in a number of ways

Unsupervised soft clustering

- Fuzzy c-means
- Rough k-means
- Evidential c-means (belief functions)
- Possibilistic clustering
- Hybridizations of the above

Relationships between soft clustering

- Principles of rough k-means have been used to create rough set representations of fuzzy c-means clustering
- Useful since we cannot show the actual cluster set after fuzzy c-means
- Similar work can be done with evidential c-means

Cluster quality indices

- Distance based for fuzzy, rough, and evidential
- Cost benefit analysis based on decision theoretic rough sets
- Study the cluster quality after granular transformation

Dynamic Unsupervised learning

- Updating clusters
- Merging clusters
- Deleting clusters
- Splitting clusters
- Use the above for distributed clustering