Name: Waseem Ahmed
Supervisor: Prof. Tanvir Ahmad
Co-supervisor: Prof. M. M. Sufyan Beg
Department: Computer Engineering
Title of Thesis: Intercommunication amongst various Information Systems applying Fuzzy Rough Sets

Abstract

Rough Set Theory (RST) has been an influential Soft Computing tool for vague and imperfect data analysis in the area of Artificial Intelligence. The essential usage of RST is to reduce the number of attributes contained in a dataset by keeping the distinguishable ability of the original system.

Initially, the RST based data analysis was applied to single-valued information systems. Later, it was found that in many cases, some attribute values are set-values (multi-values). Information Systems (IS) that follow such conditions are termed as Set-valued Information System (SIS). Attribute values in SIS are crisp values. However, in real situations, these crisp set-values for some attributes could be real or fuzzy values. The method to study SIS with discrete set-valued attributes cannot directly handle fuzzy set-valued attributes.

Therefore, the existing fuzzy rough model for handling Information system with fuzzy set-values needs some changes. In this work, a novel Fuzzy Set-valued Information System (FSIS) is proposed and thereafter, we focus on the introduction of fuzzy similarity relation for FSIS, considering similarity between the objects, which is basically a generalization of the fuzzy similarity relation for SIS.

Based on the fuzzy similarity relation, the concept of reduct, superfluous and discernibility matrix is studied. Simplest way of finding the reduct and relative reduct is to use the concept of discernibility matrix and Boolean reasoning technique. However, in this approach, every pair of

objects must be compared to each other, for constructing discernibility matrix. Hence such algorithms consume more time in handling large datasets.

In this work, information entropy based method is used, to find the reduct and relative reduct for FSIS under the framework of the FRS model. Thereafter, the significance of a candidate attribute is measured, using conditional and relative entropy, based on which, three greedy forward algorithms are discussed, to find attribute reduct for FSIS.

However, existing Fuzzy Rough Set (FRS) model for SIS suffer from some limitation, i.e., it does not fit a dataset well and can lead to samples misclassification, due to existence of large overlap between samples. So a fitting FRS model is introduced for handling FSIS. This approach can fit FSIS well and prevent samples misclassification. Later, a greedy forward algorithm for finding reduct for FSIS is constructed.

Another problem encountered in Information system is the communication between Information Systems (IS). This work aims to discuss important properties related to communication between FSIS. Fuzzy relation mapping from the perspective of FSIS is discussed. Homomorphism mapping plays a vital role in communication between FSIS. In this work, the concept of homomorphism for FSIS is explored and important properties related to communication between FSIS under FRS model are discussed. It is proved that feature selection and other properties of the original FSIS and the corresponding image FSIS are assured under consistent and compatible homomorphism.

An experiment was conducted on a sample population of the real dataset and the classification accuracy of proposed FSIS is compared with existing FIS and SIS using existing feature selection algorithms and it is found that the proposed FSIS is more effective than existing IS.