Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Pattern Recognition Letters 26 (2005) 965–975

Pattern Recognition Letters

www.elsevier.com/locate/patrec

On fuzzy-rough sets approach to feature selection

Rajen B. Bhatt *, M. Gopal

Control Group, Department of Electrical Engineering, Indian Institute of Technology Delhi, Hauz Khas, New Delhi - 110 016, India

Received 8 September 2004
Available online 25 November 2004

Abstract

In this paper, we have shown that the fuzzy-rough set attribute reduction algorithm [Jenson, R., Shen, Q., 2002. Fuzzy-rough sets for descriptive dimensionality reduction. In: Proceedings of IEEE International Conference on Fuzzy Systems, FUZZ-IEEE'02, May 12–17, pp. 29–34] is not convergent on many real datasets due to its poorly designed termination criteria; and the computational complexity of the algorithm increases exponentially with increase in the number of input variables and in multiplication with the size of data patterns. Based on natural properties of fuzzy *t*-norm and *t*-conorm, we have put forward the concept of fuzzy-rough sets on compact computational domain, which is then utilized to improve the computational efficiency of FRSAR algorithm. Speed up factor as high as 622 have been achieved with this concept with improved accuracy. We also restructure the algorithm with efficient termination criteria to achieve the convergence on all the datasets and to improve the reliability of selected set of features.

Keywords: Computational complexity; Feature selection; Fuzzy sets; Rough sets

1. Introduction

One aspect common to pattern recognition problems is feature dimensionality reduction. Two different approaches for feature dimensionality reduction are feature selection and feature extraction. The former is to find a set of input variables, from given set of input candidates, which really affect the output. The later reduces the original set of features into a linearly or nonlinearly transformed set. Feature selection allows to discard some of the irrelevant features and better performance may be achieved by discarding such features (Fukunaga, 1972; Mucciardi and Gose, 1971; Steppe et al., 1996).

In their paper (Jenson and Shen, 2002) presented an approach to dimensionality reduction by applying the concept of fuzzy-rough sets. Indeed the concept is very interesting and systematic one,

* Corresponding author. Tel.: +91 11 26596133.
E-mail addresses: bhattrajen@ee.iitd.ernet.in (R.B. Bhatt), mgopal@ee.iitd.ernet.in (M. Gopal).

0167-8655/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2004.09.044

since by considering the degree of belongingness of real-valued data to the fuzzy sets, important information loss may be recovered. Fuzzy-rough dependency measure is natural extension to Pawlak's measure for dependency degree (Pawlak, 1991). Also, feature selection stage is separated from model building stage. This approach allows building a fuzzy model directly on reduced dimension such as given by (Basak et al., 1998; De et al., 1999), and once the model has been developed, it can be tuned further by neuro-fuzzy or genetic optimization technique to improve classification/regression accuracy. Further, algorithm does not require supplying any external parameter or terminating criteria from the user.

However, the algorithm has not been applied to various benchmark problems. The purpose here is to describe our experience in applying the FRSAR (fuzzy rough set attribute reduction) algorithm, whose performance, we found, is severely compromised due to two flaws:

1. As per our investigation, this algorithm is not convergent (i.e., may be trapped into an infinite search loop) on many real data sets due to its poorly designed termination criteria and it may select unreliable features.
2. The computational complexity of the algorithm increases exponentially with Cartesian product of the input frame of cognition (i.e. Cartesian product of collection of fuzzy sets on each input attribute) and in multiplication with size of the data patterns. This makes the algorithm almost infeasible on large dimensional problems.

The main contribution of this paper is to put forward the concept of fuzzy-rough sets on compact computational domain based on the properties of fuzzy t-norm and t-conorm operators. This modified definition has been utilized to speed up existing FRSAR algorithm. We also restructure the termination criteria of the algorithm to make it convergent on subset of features within finite time and to reduce the number of levels and nodes required to evaluate in search tree.

This paper is organized as follows: Section 2 discusses the fuzzy-rough sets and FRSAR. In Section 3, we put forward the concept of fuzzy-rough sets on compact computational domain. Based on this definition, Section 4 builds improved feature selection algorithm, which is then compared with FRSAR. Experimental results, comparison with results available from the literature, and discussions are given in Section 5. Section 6 concludes the paper.

2. Theoretical foundations

A typical fuzzy classification/regression problem can be described as follows. A universe of objects or set of learning patterns $U = \{x_i | i = 1, \ldots, n\}$ are described by a collection of input fuzzy attributes $\{P_1, P_2, \ldots, P_p\}$. Each attribute measures some important feature of an object and is limited to usually a small set of fuzzy linguistic terms $A(P_i) = \{F_{ik} | k = 1, \ldots, C_i\}$. Each object $x_i \in U$ is classified by a set of classes $A(Q) = \{F_l | l = 1, \ldots, C_Q\}$.

Each $F_l \in A(Q)$ may be a crisp or fuzzy set and Q is decision attribute.

The set $U/P = \{F_{ik} | i = 1, \ldots, p;\ k = 1, \ldots, C_i\}$ can be regarded as fuzzy partitions of U by a set of attributes P using fuzzy similarity relations R on U (Radzikowska and Kerre, 2002).

The set U/P builds basic partitions about the domain of interest. Given any arbitrary fuzzy set A on the universe U, we can approximate it using basic fuzzy partitions. However, due to limited discernibility of objects, fuzzy set approximation may not be perfect. Fuzzy-rough sets are a pair of approximation degrees $\langle \underline{\mu}, \bar{\mu} \rangle$ with which we can 'certainly' and 'possibly' approximate arbitrary fuzzy set A using single or composite fuzzy set from basic fuzzy partitions.

More formal definition of fuzzy-rough sets is given below.

Definition 1. Given arbitrary fuzzy set $\mu_A(x) : U \to [0, 1]$; $\forall x \in U$ and $F_{ik} \in U/P$. According to (Radzikowska and Kerre, 2002) the fuzzy-rough set is a tuple $\langle \mu_{\underline{A}}, \mu_{\overline{A}} \rangle$, where lower and upper approximation membership functions are defined by