Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 96 (2016) 403 – 409

Procedia Computer Science

20th International Conference on Knowledge Based and Intelligent Information and Engineering Systems, KES2016, 5-7 September 2016, York, United Kingdom

An Innovative Approach for Attribute Reduction using Rough Sets and Flower Pollination Optimisation

Waleed Yamany[a,*], Eid Emary[b], Aboul Ella Hassanien[b], Gerald Schaefer[c], Shao Ying Zhu[d]

[a]*Faculty of Computers and Information, Fayoum University, Egypt*
[b]*Faculty of Computers & Information, Cairo University, Egypt*
[c]*Department of Computer Science, Loughborough University, U.K.*
[d]*Department of Computing and Mathematics, University of Derby, U.K.*

Abstract

Optimal search is a major challenge for wrapper-based attribute reduction. Rough sets have been used with much success, but current hill-climbing rough set approaches to attribute reduction are insufficient for finding optimal solutions. In this paper, we propose an innovative use of an intelligent optimisation method, namely the flower search algorithm (FSA), with rough sets for attribute reduction. FSA is a relatively recent computational intelligence algorithm, which is inspired by the pollination process of flowers. For many applications, the attribute space, besides being very large, is also rough with many different local minima which makes it difficult to converge towards an optimal solution. FSA can adaptively search the attribute space for optimal attribute combinations that maximise a given fitness function, with the fitness function used in our work being rough set-based classification. Experimental results on various benchmark datasets from the UCI repository confirm our technique to perform well in comparison with competing methods.

Peer-review under responsibility of KES International

Keywords: Pattern recognition; attribute reduction; rough sets; flower pollination optimisation.

1. Introduction

Attribute reduction is one of the essential problems in the fields of data mining, machine learning, and pattern recognition[1]. Attribute reduction is mainly concerned with selecting the smallest subset of attributes for a given problem while preserving a suitably high accuracy in representing the original attributes[2]. In real world problems, attribute reduction is often a necessity due to the presence of noisy, misleading or irrelevant attributes[3], while attribute reduction enables data processing techniques such as machine learning algorithms to yield good performance[2].

* Corresponding author.
E-mail address: wsy00@fayoum.edu.eg

Peer-review under responsibility of KES International
doi:10.1016/j.procs.2016.08.083

One of the well-known approaches for attribute reduction is based on rough set theory[4] which has the inherent ability to deal with vagueness and uncertainty in data analysis. Roughs sets have been extensively used in data mining[5], machine learning[6] and other fields. Here, knowledge is considered as a kind of discriminability, which can also be employed as an instrument to reduce attribute dimensionality and establish data dependencies. While various rough set-based algorithms for attribute reduction have been proposed, the main idea of these methods is to detect minimal reducts by creating every conceivable reduct and subsequently selecting the one with smallest length. This can be performed by developing a kind of detectability capacity from a given dataset and then streamlining it[6]. On the other hand, the number of possible subsets is typically very large and considering all attribute subsets for choosing the optimal one is considered an NP-hard problem. To address this, rough sets can be integrated with optimisation approaches such as genetic algorithms[14], ant colony optimisation[8], or particle swarm optimisation[9].

Wroblewski[7] integrated a genetic algorithm (GA) with a greedy algorithm to produce small reducts, however could not demonstrate that the produced subset is a reduct. ElAlami[13] also made use of GAs to locate ideal relevant attributes. Zhai et al.[14] proposed an incorporated attribute extraction approach that uses both rough sets and GAs. Jensen and Shen[8] applied ant colony optimisation to find rough set reducts, while Wang et al.[9] developed a method for attribute reduction using particle swarm optimisation hybridised with rough sets. Although, stochastic techniques can yield strong solutions for global optimisation, this is accomplished at the expense of computational cost[2].

In this paper, we propose a new attribute reduction technique. In our approach, we make use of the Flower Search Algorithm (FSA) to discover ideal attribute subsets (reducts). FSA is a relatively new evolutionary computation algorithm proposed by Yang[10] and is based on the flower pollination process of flowering plants. FSA can adaptively search the attribute space for optimal attribute combinations that maximise a given fitness function, with the fitness function used in our work being rough set-based classification. Experimental results on various benchmark datasets from the UCI repository confirm our technique to perform well in comparison with competing methods.

The organisation of the remainder of the paper is as follows: Section 2 describes the fundamentals of rough set theory, while Section 3 explains the flower search algorithm. Our FSA algorithm based on rough sets for attribute reduction (FLRSAR) is then presented in Section 4. Experimental results are given in Section 5, while Section 6 concludes the paper.

2. Rough Set Theory

In this section, we present some of the necessary fundamentals for rough set (RS) theory and RS-based feature selection. For more exhaustive descriptions of the theory, we refer to[4] and other publications.

Let $I = (O, S, B, f)$ be an information system, such that O is a finite non-empty set of instances, S is a finite non-empty set of attributes, and B is the set union of attribute scopes such that $B = \bigcup_{s \in S} B_s$ for B_s indicate the value scope of attribute s. $f : O \times S \longrightarrow B$ is an information function that associates a unique magnitude of each attribute with every instance in O such that $f(x, s) \in B_s$ for any $s \in S$ and $x \in O$.

For any $P \subseteq S$ there exists an associated indiscernibility relationship $IND(Z)$

$$IND(Z) = \{(x, y) \in O \times O | \forall s \in p, f(x, s) = f(y, s)\}. \tag{1}$$

The partition of O, deduced by $IND(Z)$ is indicated by $O/IND(Z)$ and can be computed as

$$O/IND(Z) = \otimes\{s \in Z : O/IND(\{s\})\}, \tag{2}$$

where

$$R \otimes K = \{X \bigcap Y : \forall R \in X, \forall K \in Y, X \bigcap Y \neq \emptyset\}. \tag{3}$$

For subset $X \subseteq O$ and equivalence relationship $IND(Z)$, the Z-lower and Z-upper approximations of X are determined as

$$Z_*(X) = \{x \in O : [x]_Z \subseteq X\} \tag{4}$$

and

$$Z^*(X) = \{x \in O : [x]_Z \bigcap X \neq \emptyset\}, \tag{5}$$

respectively.