

Accepted Manuscript

Title: A Parallel Rough Set based Dependency Calculation Method for Efficient Feature Selection

Authors: Muhammad Summair Raza, Usman Qamar

PII: S1568-4946(17)30609-9
DOI: <https://doi.org/10.1016/j.asoc.2017.10.006>
Reference: ASOC 4501

To appear in: *Applied Soft Computing*

Received date: 4-8-2016
Revised date: 27-9-2017
Accepted date: 1-10-2017



Please cite this article as: Muhammad Summair Raza, Usman Qamar, A Parallel Rough Set based Dependency Calculation Method for Efficient Feature Selection, Applied Soft Computing Journal <https://doi.org/10.1016/j.asoc.2017.10.006>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

A Parallel Rough Set based Dependency Calculation Method for Efficient Feature Selection

Muhammad Summair Raza¹, Usman Qamar²

^{1,2} Computer Engineering Department, College of Electrical and Mechanical Engineering
National University of Sciences and Technology (NUST), Islamabad, Pakistan
{¹summair.raza, ²usmanq}@ceme.nust.edu.pk

Highlights

- Rough-sets feature selection is dependent upon positive region.
- The proposed method introduces a new algorithm for parallel dependency calculation.
- This allows us to avoid the computation exhaustive positive region.
- Results shown an average of 68% reduction in computation time.
- 96% decrease in required runtime memory is also noted.

Abstract

Feature selection is the process of selecting a subset of features without losing significant amount of useful information that is available in the original dataset. The use of rough set theory remains a prominent tool for this purpose. A number of techniques based on rough sets are based on the use of positive region based dependency measure which is a computationally expensive task. In this paper, we propose a novel dependency calculation technique termed as parallel dependency calculation technique. We propose to calculate dependency by directly finding the positive region based objects without calculating the positive region itself. As the objects are independent of each other so we propose to search these objects in parallel. The proposed technique was tested against the conventional dependency calculation technique and the results showed significant increase in performance, that is, overall 63.7% reduction in execution time and 96% reduction in required runtime memory was observed. In case of feature selection algorithms, there was 68% reduction in execution time when the proposed dependency calculation technique was used. The proposed method not only provides accuracy but is computationally less demanding than the conventional positive region based approach.

Download English Version:

<https://daneshyari.com/en/article/11002690>

Download Persian Version:

<https://daneshyari.com/article/11002690>

[Daneshyari.com](https://daneshyari.com)