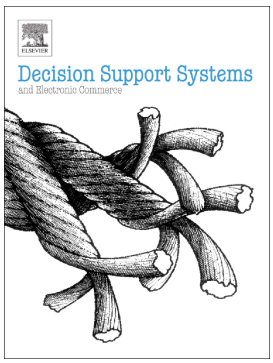

Feature selection using firefly optimization for classification and regression models

Li Zhang, Kamlesh Mistry, Chee Peng Lim, Siew Chin Neoh

PII: S0167-9236(17)30220-8
DOI: doi:10.1016/j.dss.2017.12.001
Reference: DECSUP 12902

To appear in: *Decision Support Systems*

Received date: 16 April 2017
Revised date: 19 November 2017
Accepted date: 3 December 2017

Please cite this article as: Li Zhang, Kamlesh Mistry, Chee Peng Lim, Siew Chin Neoh , Feature selection using firefly optimization for classification and regression models. The address for the corresponding author was captured as affiliation for all authors. Please check if appropriate. Decsup(2017), doi:10.1016/j.dss.2017.12.001

Feature Selection Using Firefly Optimization for Classification and Regression Models

Li Zhang[1], Kamlesh Mistry[2], Chee Peng Lim[3] and Siew Chin Neoh[4]

[1]Computational Intelligence Research Group
Department of Computer Science and Digital Technologies
Faculty of Engineering and Environment
University of Northumbria
Newcastle, UK, NE1 8ST

[2]School of Computing
Teesside University
TS1 3BA, UK

[3]Institute for Intelligent Systems Research and Innovation
Deakin University
Waurn Ponds, VIC 3216, Australia

[4]Faculty of Engineering, Technology and Built Environment
UCSI University
Malaysia

Email: li.zhang@northumbria.ac.uk; k.mistry@tees.ac.uk; chee.lim@deakin.edu.au; u_jane80@yahoo.co.uk

Abstract.

In this research, we propose a variant of the Firefly Algorithm (FA) for discriminative feature selection in classification and regression models for supporting decision making processes using data-based learning methods. The FA variant employs Simulated Annealing (SA)-enhanced local and global promising solutions, chaotic-accelerated attractiveness parameters and diversion mechanisms of weak solutions to escape from the local optimum trap and mitigate the premature convergence problem in the original FA algorithm. A total of 29 classification and 11 regression benchmark data sets have been used to evaluate the efficiency of the proposed FA model. It shows statistically significant improvements over other state-of-the-art FA variants and classical search methods for diverse feature selection problems. In short, the proposed FA variant offers an effective method to