

Chapter 17

Particle Swarm Optimization Algorithm and its Hybrid Variants for Feature Subset Selection

Basabi Chakraborty
Iwate Prefectural University, Japan

ABSTRACT

Selecting an optimum subset of features from a large set of features is an important pre-processing step for pattern classification, data mining, or machine learning applications. Feature subset selection basically comprises of defining a criterion function for evaluation of the feature subset and developing a search strategy to find the best feature subset from a large number of feature subsets. Lots of mathematical and statistical techniques have been proposed so far. Recently biologically inspired computing is gaining popularity for solving real world problems for their more flexibility compared to traditional statistical or mathematical techniques. In this chapter, the role of Particle Swarm Optimization (PSO), one of the recently developed bio-inspired evolutionary computational (EC) approaches in designing algorithms for producing optimal feature subset from a large feature set, is examined. A state of the art review on Particle Swarm Optimization algorithms and its hybrids with other soft computing techniques for feature subset selection are presented followed by author's proposals of PSO based algorithms. Simple simulation experiments with benchmark data sets and their results are shown to evaluate their respective effectiveness and comparative performance in selecting best feature subset from a set of features.

DOI: 10.4018/978-1-4666-2518-1.ch017

INTRODUCTION

Selecting an optimal subset from a large set of features is of prime importance in pattern classification, machine learning and data mining applications. The main objective of feature selection is to retain the optimum discriminatory information necessary for the classification or knowledge extraction process and to reduce the dimensionality of the measurement space so that effective and easily computable algorithms for the next step of analysis of the data can be devised. To facilitate the selection process, the quality of any feature or a set of features has to be evaluated via some well designed criterion function. More important task is to find out the best feature subset from a large number of subsets as it is well known that the best two individual features may not comprise the best feature subset of two features. However, an exhaustive search is computationally infeasible in practice giving rise to the need of developing efficient optimal search strategies. The research on feature subset selection or dimensionality reduction has a long history and well known methodologies based on mathematics and statistics are available in (Bishop, 2006; Devijver, 1982; Duda, 2001).

General feature selection techniques are based on defining a criterion function to evaluate a feature or a set of features and developing a search strategy to find out the best feature subset according to the defined criterion function. The algorithms are classified as wrapper algorithms (John, 1994) if the evaluation of the feature subset involves the classifier that is to be used in the next stage while the algorithms that evaluate the features by their intrinsic discriminatory property irrespective of the classifier, are known as filter algorithms (Almuallim, 2006; Kira, 1992). Both approaches have relative merits and demerits. Wrapper approaches can yield better performance but they incur high computational cost while filter approaches are computationally light with lower performance. Recently hybrid approaches are also

being proposed to achieve a good balance between the computational efficiency of filter models and the performance accuracy of the wrapper models. Lot of criterion functions for evaluation of feature set as well as several search strategies for finding the best feature subset have been developed so far. Traditional statistical or mathematical techniques are computationally heavy, especially for large feature space and complexity of computation increases with increasing number of features. The problem of finding out the best feature subset from a large set of features is basically a combinatorial optimization problem and heuristic search techniques are designed for optimal or sub optimal solutions compromising computational time and effectiveness of the selected feature subset.

Recently bio-inspired computational approaches (Olariu, 2006; Yun, 2011) are gaining popularity for solving real world optimization problems for their more flexibility and better computational ability to work in an environment of uncertainty, imprecision and implicit knowledge. Artificial Neural Network (ANN), the most popular computational paradigm mimicking human brain, is an active research area and has numerous applications in pattern classification (Bishop, 1995). Genetic Algorithm (GA) (Goldberg, 1989), a randomized heuristic and adaptive search technique based on the principal of natural selection, is the most widely used evolutionary algorithm for solving optimization problems where the search space is large. Several approaches for solving feature subset selection problem with GA can be found in the literature (Siedlecki, 1989; Yang, 1998). Particle Swarm Optimization (PSO) proposed by Kennedy and Eberhart (1995), a member of evolutionary computational algorithms, has been exploited recently for optimization problems including feature subset selection (Liu, 2005). Ant Colony Optimization (ACO) (Dorigo, 2001), another popular bio inspired algorithm and one of the best known examples of swarm intelligence systems, is a meta heuristic inspired by the behaviour of real ants in their search for the shortest paths to food

16 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/particle-swarm-optimization-algorithm-its/72503

Related Content

A New Fuzzy Rule Interpolation Approach to Terrorism Risk Assessment

Shangzhu Jin, Jike Geand Jun Peng (2017). *International Journal of Software Science and Computational Intelligence* (pp. 16-36).

www.irma-international.org/article/a-new-fuzzy-rule-interpolation-approach-to-terrorism-risk-assessment/190316

Cognitive Theme Preserving Color Transfer for Fabric Design

Dejun Zheng, George Baci, Yu Hanand Jinlian Hu (2012). *International Journal of Software Science and Computational Intelligence* (pp. 38-61).

www.irma-international.org/article/cognitive-theme-preserving-color-transfer/76269

Simultaneous Tolerance Synthesis for Manufacturing and Quality using Evolutionary Algorithms

Y. S. Rao, C. S. P. Rao, G. Ranga Janardhanaand Pandu R. Vundavilli (2013). *Modeling Applications and Theoretical Innovations in Interdisciplinary Evolutionary Computation* (pp. 54-72).

www.irma-international.org/chapter/simultaneous-tolerance-synthesis-manufacturing-quality/74923

Online Approaches to Missing Data Estimation

Tshilidzi Marwala (2009). *Computational Intelligence for Missing Data Imputation, Estimation, and Management: Knowledge Optimization Techniques* (pp. 165-186).

www.irma-international.org/chapter/online-approaches-missing-data-estimation/6800

Subspace Clustering of High Dimensional Data Using Differential Evolution

Parul Agarwaland Shikha Mehta (2019). *Nature-Inspired Algorithms for Big Data Frameworks* (pp. 47-74).

www.irma-international.org/chapter/subspace-clustering-of-high-dimensional-data-using-differential-evolution/213030