

Chapter 30

A Hybrid GSA–K–Mean Classifier Algorithm to Predict Diabetes Mellitus

Rojalina Priyadarshini

School of Computer Science & Engineering, KIIT University, Bhubaneswar, India

Rabindra Kumar Barik

School of Computer Application, KIIT University, Bhubaneswar, India

Nilamadhab Dash

Department of Information Technology, C.V. Raman College of Engineering, Bhubaneswar, India

Brojo Kishore Mishra

 <https://orcid.org/0000-0002-7836-052X>

Department of Information Technology, C. V. Raman College of Engineering, Bhubaneswar, India

Rachita Misra

Department of Information Technology, C. V. Raman College of Engineering, Bhubaneswar, India

ABSTRACT

Lots of research has been carried out globally to design a machine classifier which could predict it from some physical and bio-medical parameters. In this work a hybrid machine learning classifier has been proposed to design an artificial predictor to correctly classify diabetic and non-diabetic people. The classifier is an amalgamation of the widely used K-means algorithm and Gravitational search algorithm (GSA). GSA has been used as an optimization tool which will compute the best centroids from the two classes of training data; the positive class (who are diabetic) and negative class (who are non-diabetic). In K-means algorithm instead of using random samples as initial cluster head, the optimized centroids from GSA are used as the cluster centers. The inherent problem associated with k-means algorithm is the initial placement of cluster centers, which may cause convergence delay thereby degrading the overall performance. This problem is tried to overcome by using a combined GSA and K-means.

DOI: 10.4018/978-1-7998-2460-2.ch030

1. INTRODUCTION

Diabetes mellitus is a chronic disease, whose root cause is insufficient production of insulin in a patient's body. It is a type of metabolic diseases differentiated by high blood sugar (glucose) levels that result from flaws in insulin hormone emission, or its action or both. Three types of Diabetes Mellitus are found and it is being recorded from a study done by Public Health Foundation in India gives the information that nearly 44lakh Indians between the age group from 20 to 79 years is not aware of the fact that they are suffering from Diabetes. The statistics strained by International Diabetes Foundation informs that, about 50 million Diabetic patients exist in India (Alice & Balachandran, 2015). Diabetes is a serious disease that reduces the level of insulin which helps to communicate glucose into the blood platelets. As a result, some serious difficulties may arise in the human body and may lead to stroke, heart disease, kidney failure, retinopathy, paralysis and nephropathy by which the vision of a patient is affected. The consequences of diabetes are loss of weight, obscured vision, infections, frequent urination etc.

Experimental methods have proved to be complex and expensive and time consuming for this work. So now days different soft computing approaches are used for this work (Nurhayati et al.). In the past, a lot many heuristic optimization algorithms (Das et al., 2011; Geem et al., 2001; Yang et al., 2009) and machine learning approaches are widely used for diabetes detection (Sudharsan et al., 2015). The learning and training in machine learning techniques can be broadly classified into two basic types; supervised and unsupervised. In supervised learning the output is priory known to the network. Whereas, in an unsupervised learning' the output is not known beforehand. Both supervised and unsupervised algorithms are being extensively experimented to accomplish the same task. The Artificial Neural Network (ANN) (Scott et al., 2008) Support Vector Machine (SVM) (Vijayan and Anjali, 2015), and Extreme Learning Machine (ELM) are being used by many researchers for the same problem. But all these methods have their inherent disadvantages. High Time Complexity, slow convergence, getting stuck into local optima, are some inherent problems associated with the classical techniques (h. Navarro et al.(2014)). To improve the efficiency of classical methods, hybrid algorithms are used to avoid the limitations of individual algorithms used in isolation.

Gravitational Search algorithm (Rashedi et al., 2009) is a recent algorithm that has been motivated by the Newtonian's law of gravity and motion. GSA has already been explored in many areas and found to be efficient in various applications (Mohd Sabri et al., 2013, Eldos & Al Qasim, 2009). At present, there are various variants of GSA (Precup, 2012; Rashedi et al., 2010; Purcaru, 2013), which have been developed to enhance and improve the original version.

In this paper, basic GSA is used along with k-means algorithm, which is a simple and powerful clustering algorithm to determine the classes of people whether they are diabetic or not. The proposed work uses the unsupervised learning model to accomplish the work. Many existing optimization algorithms are present in current scenario, that are used for doing classification and prediction tasks in different domains. The examples could be ant colony optimization (ACO) for classifying gene expression micro array data (Schaefer et al., 2016) and for automation of blood vessel segmentation used in retinal diagnostic system (Asad et al., 2014)), Gaussian kernels for texture detection (Hemalatha et al., (2016)), Fire fly algorithm for designing low noise amplifier (R. Kumar et al. (2016), and for obtaining optimized scaling factor to be embedded for medical data exchange (Dey et al., 2014), Artificial immune system for anomaly detection in the area of ambient assisted living(Bersch, 2013), fuzzy logic based ACO to calculate trust value of different entities (Sarkar et al., 2015), cuckoo search to find the best qualitative

13 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/a-hybrid-gsa-k-mean-classifier-algorithm-to-predict-diabetes-mellitus/252045

Related Content

Brave New Digital Tools for Action Research in Education: A Beginner's Guide

Reinhard Bauer, Klaus Himpsl-Gutermann, Martin Sankofi, Petra Szucsichand Ruth Petz (2020). *Cognitive Analytics: Concepts, Methodologies, Tools, and Applications* (pp. 469-490).

www.irma-international.org/chapter/brave-new-digital-tools-for-action-research-in-education/252039

Data Accuracy Considerations with mHealth

Zaid Zekiria Sako, Vass Karpathiou, Sasan Adibiand Nilmini Wickramasinghe (2020). *Cognitive Analytics: Concepts, Methodologies, Tools, and Applications* (pp. 1623-1638).

www.irma-international.org/chapter/data-accuracy-considerations-with-mhealth/252102

Puzzling the Picture using Grounded Theory

Elisabeth E. Bennett (2020). *Cognitive Analytics: Concepts, Methodologies, Tools, and Applications* (pp. 1778-1792).

www.irma-international.org/chapter/puzzling-the-picture-using-grounded-theory/252111

Music Emotions Recognition by Machine Learning With Cognitive Classification Methodologies

Junjie Bai, Kan Luo, Jun Peng, Jinliang Shi, Ying Wu, Lixiao Feng, Jianqing Liand Yingxu Wang (2020). *Cognitive Analytics: Concepts, Methodologies, Tools, and Applications* (pp. 1028-1041).

www.irma-international.org/chapter/music-emotions-recognition-by-machine-learning-with-cognitive-classification-methodologies/252068

Jung's Collective Unconscious, Integrative (Mind-Body-Spirit) Yoga, and Self-Realization

Manoj Sharma (2019). *Media Models to Foster Collective Human Coherence in the PSYCHecology* (pp. 93-108).

www.irma-international.org/chapter/jungs-collective-unconscious-integrative-mind-body-spirit-yoga-and-self-realization/229330