Chapter 55
Optimal Fuzzy Cluster Partitioning by Crow Search Meta-Heuristic for Biomedical Data Analysis

Janmenjoy Nayak
https://orcid.org/0000-0002-9746-6557
Aditya Institute of Technology and Management, India

Pandit Byomakesha Dash
https://orcid.org/0000-0003-1643-9651
Veer Surendra Sai University of Technology, Burla, India

Bighnaraj Naik
https://orcid.org/0000-0002-9761-8389
Veer Surendra Sai University of Technology (VSSUT), Odisha, India

Danilo Pelusi
https://orcid.org/0000-0003-0889-278X
University of Teramo, Italy

ABSTRACT

Biomedical data is often more unstructured in nature, and biomedical data processing task is becoming more complex day by day. Thus, biomedical informatics requires competent data analysis and data mining techniques for designing decision support system's framework to solve clinical and heathcare-related issues. Due to increasingly large and complex data sets and demand of biomedical informatics research, researchers are attracted towards automated machine learning models. This paper is proposed to design an efficient machine learning model based on fuzzy c-means with meta-heuristic optimizations for biomedical data analysis and clustering. The main contributions of this paper are 1) projecting an efficient machine learning model based on fuzzy c-means and meta-heuristic optimization for biomedical data classification, 2) employing benchmark validation techniques and critical hypothesises testing, and 3) providing a background for biomedical data processing with a view of data processing and mining.

DOI: 10.4018/979-8-3693-3026-5.ch055

1. INTRODUCTION

Data mining is a fairly novel concept that appeared as a new approach to analysis of data and knowledge discovery. It is a multidisciplinary field which includes the assemblage of various developments such as pattern recognition, neural network, database design, artificial intelligence, data visualization, machine learning and many others. It has been appeared as a basic research area with progressively more significant application in research, businesses, medicine, education, science, engineering and one of the best inspiring fields of research that is become progressively more admired in medical field. There are different systematic tools based on soft computing, statistics etc., are used for the appropriate data analysis. In the previous days, in order to extract the useful information, the investigators have utilized recorded data. But in current days it is quite contrast because of increment of data amount as well as exponential development of both science and technology (Li et al., 1998). Particularly, in the medical field, data are inundated with a number of features and unique characteristics which may be important at some of the later stage of diseases diagnosis. Therefore, some automated tools need to be developed for dealing with some of the common as well as complex problems in medical health care system or biomedical applications. In present days, data mining is becoming famous in medical field because there is a requirement of well-organized investigative methodology for identifying important and indefinite information in health data (Mitra, et al., 2002). In medical field, data mining offers various profits such as detection of medical treatment processes, recognition of the scam in health security, accessibility of medical solution to the patients at lesser cost and finding of reasons of diseases. Mining techniques also assist to the researchers in medical field for accomplishing competent policies of healthcare, rising health profiles of individuals, creating drug approval schemes etc. Data produced by the medical organizations is extremely huge and multifaceted due to which it is hard to examine the information in order to make significant conclusion regarding the health of a patient. This information holds particulars concerning hospitals, patients, medical claims, cost of treatments etc. So, there is a demand to generate a commanding tool for examining and removing vital information from this difficult data. The results of data mining technologies are to present profits to medical fields for assembling the patients having same type of diseases or fitness problems, so that medical organization gives them effectual treatments. Modern technologies are used in medical field to improve the medical services in cost efficient way.

Biomedical data clustering is an important area of research due to the requirement of advance techniques in various biomedical applications. Many such methods have been developed to identify and resolve the issues of these applications along with some practical medical domain. Although clustering is an unsupervised method of learning, still it is considered to be one of the widest data mining approaches due to its different algorithm forms as well as suitability to fit with the problem domain. Sometimes, it is better than the supervised methods where some early training is supplied for proper functioning of the method. Clustering with biomedical data requires some specific analysis based on the biomedical scientists or doctors or field researchers. As it is common that, clustering algorithms have certain distinguished characteristics or features and those are to be suitable handle while applying to any biomedical applications. Some common features (Andreopoulos et. al., 2009) such as scalability, finding the non-required outliers, number of algorithmic dependent parameters, sensitivity to supplied inputs, nature of data type (homogeneous or heterogeneous), optimality etc. are important for analysing and applying in biomedical applications.

Data clustering is an important unsupervised method of pattern recognition, where the objects are partitioned into a set of disjoint clusters. Broadly, it may be hierarchical or partitioned. In hierarchical