ABSTRACT

Feature Selection is a data preprocessing task of selecting the more informative features from gigantic number of data. It has been an active and fruitful field of research area by a main goal of finding a feature subset that produces higher classification accuracy. In this way, it is possible to avoid excessive storage and time complexity, computational complexity and improving the results obtained by any data mining application.

Primarily in this research, the researcher focuses on Genetic Algorithm for feature selection, which is proficient, except, slow convergence. To rise above this flaw, Genetic Algorithm has been fused with local search, named as Memetic Algorithm. In this loom of hybridization, a pragmatic algorithm, baptized as Compound Featuristic Genetic Algorithm, to upliftment the Genetic Algorithm and symbolizes the best solutions.

Secondly, Class Dependent Feature Subset Selection, an algorithm is proposed namely, Core Featuristic Genetic Algorithm. Because, in a dataset, attributes are attached with class, which grasps the sticky tag of categories of the instance belongs to. To appraise the proposed algorithms, a comparative analysis is performed with the existing Genetic Algorithm.

Later the researcher deals with the recital analyses of existing and proposed feature selection algorithms which functions on heart dataset to predict the heart disease with minimum number of features.

Finally Fuzzy Classifiers, such as Fuzzy Decision Tree, Fuzzy Naive Bayes and Fuzzy Neural Networks are applied to the reduced set of the Heart dataset, obtained for classification accuracy. A comparative analysis is also exposed that Fuzzy Neural Networks with Core Featuristic Genetic Algorithm outperforms.