IJEAST

INTERNATIONAL JOURNAL
OF ENGINEERING APPLIED SCIENCE
AND TECHNOLOGY

VOLUME : 2 ISSUE : 9 Print / Issue Publication Date: 25-Mar-2018

ISSN : 2455-2143

Indexed In

Google SCHOLAR OPEN ACCESS doi Crossref

WWW.IJEAST.COM

editor@ijeast.com

International Journal of Engineering Applied Sciences and Technology, 2017
Vol. 2, Issue 9, ISSN No. 2455-2143, Pages 42-54
Published Online January 2018 in IJEAST (http://www.ijeast.com)

A REVIEW ON FEATURE SELECTION PROBLEM SOLVING USING MULTIOBJECTIVE EVOLUTIONARY OPTIMIZATION ALGORITHMS

P. D. Sheth
Ph. D. Scholar in Computer Engineering
Vishwakarma Institute of Technology, Pune

S. T. Patil
Professor in Computer Engineering
Vishwakarma Institute of Technology, Pune

***Abstract*— Feature selection plays a vital role in data mining and machine learning algorithms because of big sizes of datasets and features. Feature selection is the pre processing step where irrelevant and redundant features are removed from datasets. It would be useful in reducing dimensionality of the data, execution time and improving the predictive accuracy of classifier. Feature selection problem has number of features and classification accuracy as multiple conflicting objectives, which must be optimized simultaneously. Evolutionary computation has proven itself as effective choice to consistently reduce the number of attributes towards a better classification rate. This paper reviews important and recent algorithms published in the area of Multiobjective Evolutionary Optimization Algorithms (MOEAs) for Feature Selection. No work has reviewed papers which have been published in this area in recent five years. A detailed review containing year, MOEA, new algorithms, crossover, mutation, future scope and limitation of each algorithm is provided in chronological order.**

Keywords*— Multiobjective Evolutionary Optimization Algorithms, MOEA, Feature Selection, Evolutionary Algorithms, Machine Learning, Data Mining, Multiobjective Optimization Problem, MOO

I. INTRODUCTION

Numerous high-dimensional data mining, machine learning, pattern classifications, pattern recognition and modeling tasks often deal with the data which contain redundant, noisy, dominated or irrelevant inputs that are needed to be removed [1]. These types of data often contain raw data is generally represented as n dimensional vectors or points in an n-dimensional space. The data is frequently represented by a huge number of features. The processing of data in a high dimensional space is going to be computationally complex [2]. In such cases, the problem of the curse of dimensionality occurs when the number of features is too large than the number of available training patterns [1]. Thus, the problem of dimensionality reduction aims at reducing the number of features [2]. Most of the times, existing features are not discriminative enough. Feature selection aims to select a minimal number of features with the maximal discriminative capability. The discriminability of a feature subset requires feature subset with the high relevance to class labels. Low redundancy is aimed by the compactness within the selected feature subset [3].

Feature extraction and feature selection plays significant role in classification. They try gain a discriminative subset of features in a lower dimensional space. Both methods could be useful to find a proper feature subset, but their techniques are different. Feature extraction transforms the existing features, but only feature selection can select the appropriate feature subset in a lower dimensional space from a feature set [4]. The problem of the curse of dimensionality can be overtaken by representing the data in a lower dimensional space. The other advantages of feature selection are as follows- It lowers down the number of records required to train a classifier thus, avoids overfitting and improves the generalization performance of a classifier. The use of optimum features improves computational efficiency. Data visualization is easier and more intuitive in lower dimensional space such as 2D or 3D [3].

Feature selection is an important step in data mining tasks such as clustering, classification, regression, and time series prediction problems. It contributes in decreasing computational complexity of classifier. Initially, feature selection was defined as a single objective optimization problem. In that case, classification accuracy was only objective to be optimized. Nowadays multiobjective approaches have been presented to this problem. In addition to classification accuracy, this class of problems includes multiple objectives such as generalization capability for supervised classifiers. For unsupervised classifiers, it counterbalances the bias toward lower or higher numbers of features. In the literature preference based approaches for feature selection has seen but the necessity of finding multiple trade off solutions was clearly stated.

The multi objectives algorithm based feature selection fits properly to the problems whenever:

- Multivariate and linear methods like Principal Component Analysis (PCA) are not adequate to reduce the feature dimensions.
- The detection rate for all classes deals with mutually exclusive objectives. [3]

As per the authors' search the first paper in this area published in the year 2001 [4] and gain the publicity in the year 2010 when numbers of features in many areas become relatively large. The references [4] and [5] present recent reviews on Feature Selection using MOEA from 2000 to 2014. Some authors reviewed only the papers published on Feature Selection using Evolutionary Algorithms. No work has reviewed the papers which have been published in recent five years. Therefore this work mostly reviews the papers published since 2011 to till date. There are about 100 papers published in this area in the last 10 years.

The remainder of this paper is organized as follows. Section II describes the important facts regarding MOEA and feature selection. Section III reviews the work done on MOEAs for feature selection. Section IV discusses different aspects of MOEAs for feature selection including adopted MOEAs, chromosome representation, objective functions, evolutionary operators, selecting the final solution, performance measures. Section V discusses issues and future scope. The conclusion is given in the Section VI.

II. IMPORTANT FACTS

A. Multiobjective Optimization (MOO) Problem

Many real time problems have multiple objectives under consideration [4]. In multiobjective optimization, several competing objectives are needed to be solved simultaneously subject to certain constraints. The fundamental part in multiobjective optimization is that there is no ultimate definition of optimum. These problems produce multiple solutions. These objectives are mostly in conflict with one another. Therefore, no single solution can be found that simultaneously optimizes all objectives [6]. Multiobjective optimization problems aim to generate a set of solutions each of which is good enough to satisfy all the objectives to an extent without being dominated by any other solution in the solution space [6]. The best solution is a subjective term and depends on the requirement of the decision maker [7]. The multiobjective optimization problem can be stated as follows: [8, 9].

Minimize:

$$\vec{F}(\vec{x}) := [f_1(\vec{x}), f_2(\vec{x}), \ldots\ldots, f_k(\vec{x})] \quad (1)$$

Subject to:

$$g_i(\vec{x}) \le 0, \text{where } i = 1,2,\ldots.m \quad (2)$$

$$h_i(\vec{x}) = 0, \text{where } i = 1,2,\ldots.p \quad (3)$$

Where $\vec{x} = [x_1, x_2, \ldots\ldots x_n]^T$ the vector of decision variables is, $f_i : R^n \to R, i = 1,\ldots\ldots k$ are the objective functions and $g_i, h_j : R^n \to R, i = 1,\ldots\ldots m, j = 1,\ldots\ldots p$ are the constraint functions of the problem.

The most popular method is to generate a Pareto optimal solution set. A Pareto optimal set can be defined as a set of all Pareto optimal solutions that are not dominated by any other solution in the solution space. The respective nondominated objective solutions in the objective space are known as the Pareto front. While traversing from one Pareto solution to another, there is always a certain amount of loss with respect to a particular objective to obtain a certain amount of gain with regards to others [19]. For example, let a multi-objective minimization problem with k objectives. A feasible solution x is said to dominate another feasible solution y i.e. ($x > y$), if and only if $f_i(x) \le f_i(y)$ for $i = 1,2,3,\ldots.k$ and $f_j(x) < f_j(y)$ for at least one objective function j. The final task of multiobjective optimization is to find out solutions in the Pareto optimal set.

B. Multiobjective Evolutionary Algorithms (MOEA)

To solve the multiobjective optimization problems traditional mathematical programming based search and optimization methods such as calculus methods become difficult because their basic structure doesn't consider multiple solutions. The population-based meta-heuristic techniques such as evolutionary algorithms are well-suited for handling such conditions. Many approaches presented in [4, 7, 8, 9, 10] to solve multiobjective optimization problems.

David E. Goldberg suggested a sketch of multi-objective evolutionary algorithm (MOEA) in his book using the concept of domination in 1989. Many Evolutionary Algorithm researchers were encouraged by his book. They experimented and developed different implementations of MOEAs. MOEAs then evolved over a number of years. They have started from traditional aggregating approaches. Elitist models of Pareto-based multiobjective evolutionary algorithms were introduced in the late 1990s. Now, recently the indicator-based algorithms are proposed. The most popular Pareto-based approaches such as multiple objective GA [11], niched Pareto genetic algorithm (NPGA) [12], and non-dominated sorting genetic algorithm (NSGA) [13] were immediately tested on multiple real time applications and gain the publicity.

However, they exclude elitism and therefore cannot guarantee that the non-dominated solutions obtained during the search are preserved. The most representative elitist MOEAs which proposed then, include strength Pareto evolutionary algorithm (SPEA) [14] and SPEA2 [15], Pareto archived evolutionary strategy (PAES) [16], Pareto envelope-based selection algorithm (PESA) [17] and PESA-II [18], and non-dominated sorting genetic algorithm-II (NSGA-II) [19]. The recent applications of MOEAs for feature selection

International Journal of Engineering Applied Sciences and Technology, 2017
Vol. 2, Issue 9, ISSN No. 2455-2143, Pages 42-54
Published Online January 2018 in IJEAST (http://www.ijeast.com)

problems have used one of these Pareto-based elitist approaches. A recent trend in MOEA structure design is to implement a selection mechanism based on some performance measure, for example, the indicator-based evolutionary algorithm [20]. In recent years the NSGA-II has combined with Elitist Pareto-based Multiobjective Evolutionary Algorithm for Diversity Reinforcement (ENORA - $\mu+\lambda$) strategy. In this approach, on the basis of the nondomination level of the individual in its slot the calculation of the rank of the individuals in the population is performed.

C. Feature Selection

A framework for feature selection algorithms has been proposed by Dash and Liu [21] comprising four stages such as a generation procedure, an evaluation function, a stopping criterion, and a validation process. The generation procedure and evaluation function are the two major steps. The generation procedure is a search process that generates feature subsets for evaluation. Various searching schemes in this procedure include complete, heuristic, and random strategies. The valuation function targets to measure the differentiating ability of a feature subset for recognizing specific class labels. Evaluation functions are divided into five categories as follows: distance, information, dependence, consistency, and classifier error rate functions. The algorithms which use scores of features in different statistical tests for their correlation with the outcome variable are categorized into filter feature selection approach. The algorithms that are using subset of features and training algorithm for evaluation are called the wrapper feature selection approach. [22]

Filter methods uses feature subset according to heuristics depending on different data characteristics. Features are ranked by score considering their usefulness in discriminating classes [23] and either selected or removed from dataset. Statistical methods comprise hypothesis testing, such as Student's t-test [24, 25]. Information theory-based methods use different metrics, such as entropy, Kullback-Leibler divergence [24] or the information gain measure [26] to rank the features. A different class of filter algorithms uses a correlation based metric to measure the suitability of features. is known as the Correlation Based Feature Selection (CFS) algorithm [27, 28]. Each feature has been assessed for the ability to discriminate among them and select those individuals that best describe each class. Class labels can be predicted by computing an average score for the different dataset classes. It may result in removing some features that would be appropriate for specific class label [23].

Wrapper algorithms use an objective function that evaluates the current feature subset. This approach is classifier dependent and needs feedback in terms of classification accuracy or classification error to assess the classifier [23]. Frequently wrapper methods achieve better results, but filter methods run more efficiently. Wrapper approach still suffers from local convergence. Besides wrapper method, numerous search techniques have been applied to feature selection, such as complete search, greedy search, heuristic search, and random search [29-33]. The meta-heuristic algorithms such as genetic algorithms, ant colony optimization are high-level, efficient global search technique that provides a set of strategies to develop the heuristic algorithm. Recently researchers attempt to solve feature selection problems using meta-heuristic techniques [34].

III. LITERATURE REVIEW

Christos Emmanouilidis et. al. [36] 2001 proposed multiobjective genetic algorithms. It also includes distinct way of neural network and neurofuzzy models. The success of this work lies in the designing of feature selection as a multiobjective optimization problem based on the concept of dominance.

Gisele L. Pappa et. al. [37] 2002 addressed multiobjective attribute selection in data mining. They presented a wrapper approach based on multiobjective genetic algorithm (MOGA) to find the best subset of attributes C4.5 decision tree algorithm.

Luiz Oliveira et. al. [38] 2006 presented an ensemble feature selection approach based on a hierarchical multi-objective genetic algorithm. In the first level algorithm performs feature selection to generate a set of classifiers and then it chooses the best team of classifiers. In the supervised contexts, the problem of handwritten digit recognition with multi-layer perceptron neural networks as classifiers is solved. Where as in the unsupervised context the problem of Handwritten Month Word Recognition with Hidden Markov Model is solved.

Daniela Zaharie et. al. [39] 2007 analyzed a feature ranking technique based on the weights using MOEA. Also, they addressed the problem of comparing and aggregating different rankings. Rankings are obtained either by applying different methods to the same dataset or by applying in the context of distributed data mining tasks, the same method to different datasets.

António Gaspar Cunha et. al. [40] 2007 proposed an optimization methodology based on Reduced Pareto Set Genetic Algorithm with elitism (RPSGAe). Support Vector Machines (SVM) used as a classifier.

Kashif Waqas et. al. [41] 2009 exhibited independent subsets of features gives better accuracy. After observing the results they conclude that the approach will be tried with several variants of multi-objective genetic algorithms as well.

Bingquan Huang et. al. [42] 2009 presented a novel multiobjective feature selection approach for churn prediction in telecommunication service field. This approach is based on modified NSGA-II which selects local feature subsets followed by the nondominated solutions searching method to select the global nondominated feature subsets. The finding best solution method (FBSM) produces fitness thresholds.

This method is designed to select the global solutions where the lowest ranks are considered as the final solutions.

Ant´onio Gaspar Cunha et. al. [43] 2010 developed multi objective algorithms to maximize classifier accuracy and/or to minimize the errors produced while minimizing the number of features and similarly optimize the classifier parameters. Reduced Pareto Set Genetic Algorithm (RPSGA) is implemented for features selection and the classifier used is Support Vector Machines (SVM).

Venkatadri.M et. al. [44] 2010 presented an application of MOGA the Feature Selection problem, combining different criteria measuring the importance of the subsets of features.

Asif Ekbal et. al. [45] 2011 posed the model of finding an appropriate classifier ensemble as a multi objective optimization (MOO) problem for named entity recognition is posed. This technique follows ensembling of several classifiers those are trained using different feature representations instead of searching for the best fitting feature set for a particular classifier.

Asif Ekbal et. al. [46] 2011 developed a multiobjective technique based on simulated annealing. It aims to solve the feature selection problem in anaphora resolution.

Sujata Dash et. al. [47] 2012 applied the correlation-based feature selection method (CFS). It evaluates a subset of features by filter and wrapper approach both. In filter approach the individual's predictive ability of each feature along with the degree of redundancy among them is considered. In the wrappers approach classifiers such as J48, Random Forest and Random Trees are implemented for feature selection. The performance of classification is then evaluated by selected gene subsets.

Xiangtao Li et. al. [48] 2013 exhibited a multiobjective biogeography based optimization method. The method selects the small subset of informative gene relevant to the classification. Fisher-Markov selector chooses the 60 top gene expression data. The binary biogeography based optimization (BBBO) is proposed based on a binary migration and a binary mutation model for making biogeography based optimization suitable for the discrete problem. A multi objective binary biogeography based optimization (MOBBBO) is presented by integrating non dominated sorting and crowding distance method into the BBBO framework. In the proposed approach the MOBBBO method is used for gene selection and the the Support Vector Machine (SVM) is used as the classifier with the leave-one-out cross-validation method (LOOCV).

Bing Xue et. al. [49] 2013 introduced multi-objective particle swarm optimization (PSO) for feature selection. This method generates a Pareto front of non-dominated solutions. Two PSO-based multi-objective feature selection algorithms are investigated in this work. The first methodology utilizes the idea of non-dominated sorting into PSO to address feature selection problems. The second algorithm applies the ideas of crowding, mutation, and dominance to PSO to search for the Pareto front solutions.

Sujoy Paul et. al. [50] 2013 proposed a new decomposition based evolutionary multi-objective algorithm (MOEA/D). for feature selection and weighting. The feature vectors are selected and weighted simultaneously to project the data points to a hyper space. The data points become easier to classify by increasing the distance between data points of non-identical classes in the hyper space. MOEA/D is used to simultaneously optimize the interclass and intraclass distances with the optimal features and the scaling factor linked with them. At last, k-Nearest Neighbor (k-NN) is used to classify the data points after the reduction of the feature set.

Emiro de la Hoz et. al. [3] 2014 presented a novel multi-objective approach to an unsupervised clustering procedure based on Growing Hierarchical Self-Organising Maps (GHSOMs) approach for feature selection. It covers a new technique for unit labelling and efficient determination of the winning unit.

Bing Xue et. al. [51] 2014 constructed a multi-objective feature selection algorithm based on differential evolution (DE) approach. The multi-objective approach is compared with two conventional methods and two DE based single objective methods. The first algorithm minimizes the classification error rate where the second algorithm combines a number of features and the classification error rate into a single fitness function.

Luiz S. Oliveira et. al. [52] 2014 developed a hierarchical multi-objective genetic algorithm based ensemble feature selection scheme. At the first level of the algorithm it executes feature selection in order to generate a set of classifiers and then it chooses the best team of classifiers.

Choo Jun Tan et. al. [53] 2014 introduced a new multiobjective evolutionary algorithm based ensemble optimizer integrated with neural network models. The Modified micro Genetic Algorithm (MmGA) is used to form the ensemble optimizer.

Mrutyunjaya Panda et. al. [54] 2014 proposed the fusion of AnDE (Averaged n-Dependence Estimators) where n=1 and a variant of naive Bayes with efficient feature selection. The method is based on multi-objective evolutionary algorithm ENORA. The operations are performed with the aim of attaining a fast hybrid classifier which can efficiently learn from big data.

Christopher Smith et. al. [55] 2014 presented a hybrid multi objective evolutionary algorithm that trains and optimizes the structure of recurrent neural networks for time series prediction. They exhibited methods of selecting individual prediction models from the Pareto set of solutions. The first method selects all individuals below a threshold in the Pareto front and the second one is based on the training error. The individuals near the knee point of the Pareto front are also selected. The final method selects individuals based on the diversity of the individual predictors.

M. Anusha et. al. [56] 2014 extended k-means genetic algorithm (NLMOGA) using constraint feature selection on the selected subpopulation by maximizing the accuracy of the solution. A constrained feature selection is applied to each subpopulation by generating a new population through NLMOGA. The said method is proposed to improve the robustness of NLMOGA for different instances of MOPs. NLMOGA selects a solution from global population repository. Then the neighborhood learning promotes the evolution of each objective for the selected solution.

Nijat Mehdiyev et. al. [57] 2014 proposed a crisp and fuzzy rule-based classifiers in an ensemble model so as to derive decision rules as event patterns. The most significant feature subset is selected using a multiobjective evolutionary algorithm before implementing the ensemble classifier directly to the streaming data.

P Martín-Smith et. al. [1] 2015 has proposed a set of label-aided utility functions that make possible the impactful search of the most appropriate subset of features through an evolutionary multi-objective optimization scheme. The results from the proposed filter method exhibit less time consumption and better generalization capabilities with respect to some wrapper methods.

Zhichun Wang et. al. [22] 2015 defined a new feature redundancy measurement for estimating mutual information between features with respect to the target class (MIFS-CR). Based on a relevance measure and the new redundancy measure, Pareto based MOEA called multi-objective evolutionary algorithm using class-dependent redundancy for feature selection (MECY-FS) is presented with both the maximal relevance and the minimal redundancy.

Zhang Yong et. al. [34] 2015 stated an effective MOEA based on bare-bones particle swarm optimization incorporate two new operators. One is a reinforced memory strategy to overcome the degradation phenomenon of particles and another is hybrid mutation to improve the search ability of the proposed algorithm.

Jyoti Ahuja et. al. [35] 2015 proposed a hybrid approach. The Multi-Objective Genetic Algorithms (MOGA) at filter phase provides a non-dominated set of feature subsets. Genetic Algorithm at wrapper phase does the classifier dependent optimization. In the wrapper phase they have used support vector machine (SVM) as the classification algorithm.

A. Khan et. al. [58] 2015 projected a technique which implements NSGA – II. The fitness of a feature subset is evaluated using ID3. The testing accuracy obtained is then allocated to the fitness value in the evolution process.

Fernando Jim´enez et. al. [59] 2015 exhibited an application of classification to the data extracted from an integrated multichannel multi-skill contact canter. Evolution process is carried out using multi-objective evolutionary algorithm ENORA. They proposed an algorithm to integrate feature selection for classification, model evaluation, and decision making to choose the most suitable model in a multiobjective context according to a posterior process MOEAs for feature selection.

F. Jim´enez et. al. [60] 2016 developed a feature selection wrapper model composed by a multiobjective evolutionary algorithm, the clustering method Expectation-Maximization (EM), and the classifier C4.5. The proposed method utilized in unsupervised classification where data is extracted from a psychological test named, BASC-II (Behaviour Assessment System for Children - II). It has two objectives: Maximizing the likelihood of the clustering model and maximizing the accuracy of the obtained classifier. It will help in decision making to choose the most satisfactory model according to a posteriori process in a multiobjective context and testing.

Anita Sahoo et. al. [61] 2016 proposed two different techniques for multiobjective binary GWO algorithms. One is a scalarized approach to multi-objective Grey Wolf Optimizer (MO GWO) and the other is a Non dominated Sorting based GWO (NS GWO). These are used for wrapper based feature selection that selects optimal textural feature subset for improved classification of cervix lesions.

F. Jim´enez et. al. [62] 2016 built a regression model for online sales forecasting via a novel feature selection algorithm using multiobjective evolutionary algorithm ENORA. It is an Evolutionary NOn-dominated Radial slots based Algorithm where Random Forest algorithm is used as a classifier. The proposed model integrates feature selection for regression, model evaluation, and decision making, in order to choose the most satisfactory model.

Alejandro Rosales-P´erez et. al. [63] 2016 introduced a multiobjective evolutionary approach for data reduction. The proposed method simultaneously generates prototypes and selects features for k-NN classifiers.

Nouha Nouri et. al. [64] 2016 proposed a bi-objective blocking permutation flow shop scheduling problem. It considers the make span and total completion time as objective functions. The main interest is of this work is to propose a Genetic Algorithm based on NSGA-II for searching locally Pareto-optimal frontier for the problem. Non-dominated solutions and differences among parents are taken advantage when designing the selection operator.

Fuyu Cai et. al. [65] 2016 introduced a fuzzy criterion in multi-objective unsupervised feature selection. They applied a hybridized filter-wrapper approach (FC-MOFS). This methodology gives an efficient way to select features and to avoid misunderstanding of overlapping.

Ayas Das et. al. [2] 2017 determined the best possible constraints on the weights to be optimized. They evaluated the proposed bi-objective feature selection and weighting framework by using *k*-Nearest Neighbor (*k*-NN) classifier and the results are found quite competitive.

Yingying Zhu et. al. [4] 2017 proposed a technique for intrusion detection system which uses two approaches, The approaches are known as special domination method and predefined multiple targeted search, for population evolution. Non-dominated sorting genetic algorithm-III (NSGA-III) is used to getting a sufficient feature subset. An improved many-objective optimization algorithm (I-NSGA-III) is proposed using a novel method for niche preservation. It comprises a bias-selection process that selects the individual with the fewest selected features and a fit-selection process that selects the individual with the maximum sum weight of its objectives

Bing Xue et. al. [66] 2017 developed two multi-objective frameworks based on NSGAII and SPEA2 using filter method. Its four variants for feature selection are then constructed by applying filter based measures such as mutual information and entropy in each of the two proposed frameworks.

IV. MULTIOBJECTIVE EVOLUTIONARY OPTIMIZATION ALGORITHMS FOR FEATURE SELECTION

A. Adopted MOEAs

Different feature selection algorithms use numerous MOEAs as the underlying optimization tool. NSGA has been adopted in [13, 42]. The modified versions of NSGA, NSGA-II have been used in [1, 3, 42, 45, 51, 58, 60, 61, 65, 66] and NSGA-III has been used in [4]. A reduced Pareto set genetic algorithm (elitist) (RPSGAe) has adopted in [40] where clustering algorithm is applied to reduce the size of the Pareto optimal set. Strength Pareto Evolutionary Algorithm (SPEA) is employed in [14], and SPEA2 is adopted in [15, 22 and 66]. Multiobjective Genetic Algorithm (MOGA) is used as an underlying tool in [35-37, 44, 46, 52, 56, 61 and 64]. ENORA is applied in recent applications [54, 57, 59, 60 and 62] with improved efficiency. Bing Xue et. al. [66] 2017 adopted both SPEA2 and NSGA-II individually in their four proposed algorithms. NSGA-II is combined with MOGA in [61] and with ENORA in [60]. Zhang Yong et. al. [34] 2015 used Particle Swarm Optimization (PSO) with crowding distance as an optimization tool for feature selection of unreliable data. Jyoti Ahuja et. al. [7] 2015 proposed a hybrid method where they used MOGA in filter approach and GA in wrapper approach. PAES is applied in [16, 63] and hierarchical MOGA is applied in [38] for ensembles. Multi-objective binary biogeography based optimization (MOBBBO) is proposed and used in gene selection by Xiangtao Li et. al. [48] in 2013.

B. Chromosome representation

The first step in solving feature selection problem using MOEA is to encode a feature subset in the form of the chromosome [4]. Most of the MOEA based feature selection algorithms use binary chromosome. The length of each chromosome is taken as the total number of features d . Each bit in d can take either 1 or 0 value. If the value of the bit is 1 then corresponding feature is selected in feature subset, if the value is 0 then the corresponding feature is omitted. In [1], each individual is codified by a set of vectors, with each vector corresponding to one of the features included in the selection codified by the individual. The components of the vector correspond to the dimensions that characterize each input pattern. When dealing with large dimensional datasets, this encoding method results in very large size chromosomes. A different encoding method is taken up in [2] which use real-valued vectors to represent feature subsets. The length of chromosomes is determined by the number of selected features set in advance; each bit denotes the index of a selected feature.

C. Objective Functions

The significance of the selected features is evaluated using some classification performance metrics, which act as an objective function in regard to MOEA. One of the pioneering works in this view is [66]. In this work, two objective functions focus on the misclassification rate and the number of features. Both objective functions are aimed to minimize. Nowadays, most of the algorithms and applications proposed in this area are currently deal with objectives such as number of features and classification accuracy. When evaluating the feature subsets, the MECY-FS algorithm considers two factors: feature relevance and feature redundancy [22]. In [1] Label aided filter approach, GHSOMs [3] and FSS-MOGA [66] have used objectives namely classifier accuracy and generalization capability. BMOPSOFS [34] uses reliability, classifier accuracy, MOGA: A hybrid approach [7] uses intercorrelation, intracorrelation, an based algorithm MOEA/D [2] uses relevancy, redundancy, ENORA based FS [61] uses number of features, root mean squared error, BASC-II MOEA [60] uses likelihood of cluster, classifier accuracy, MOEA for Flow Shop Scheduling [64] uses makespan, total completion time, MOGA for Attribute Selection [37] uses classifier accuracy, computational complexity, Hybrid- MOEA for RNN [55] uses complexity, training errors, Extended NLMOGA [56] uses diversity, compactness, and MOEA/D [50] uses inter-class, intra-class distances as objectives.

Various approaches in selecting objective functions are seen in the literature although; classification performance majorly depends on the chosen classifier. Performance of the MOEA procedure may depend on the number and the kind of objective function is adopted. It is observed that comparative study with this regard is not done till date.

D. Evolutionary Operators

The evolutionary operators, crossover and mutation, are used to produce the population of the next generation in an MOEA. A crossover operator randomly selects two parents and a subset (also randomly selected) of features for each parent. These subsets are interchanged by the parents. The mutation operator applies changes to a subset of features randomly chosen among the features codified by the individual to be mutated, which is also randomly selected.

It has been noticed that one-point and uniform crossover had been the most popular choices in the reviewed references. One-point crossover has been implemented in [2-4], [20], [34-42], [44-47], [49-52] and [54-64]. The remaining MOEAs have adopted simulated binary crossover in [1], uniform crossover in [6], two point crossover in [48] and adaptive crossover in [53]. It has been reviewed that bit flip mutation is used in most of the MOEAs. Few exceptions are observed as below: some MOEAs are adopted hybrid mutation in [12], shift mutation in [48], adaptive mutation in [53], probability mutation in [58], and crowding mutation in [63].

It is observed that the exact functioning parameters and results of crossover and mutation are not described in most of the papers. It is also surveyed from available literature that most of the references have been used standard crossover and mutation operators. Very few MOEAs have been used recent operators. But, no work has found which compares standard and recent operators.

E. Selecting the final solution

In case of many existing MOEAs the stopping condition is the number of generations. As the number of generation reaches the maximum generation, MOEA stops and gives the last generation with their corresponding fitness values and fronts. As stated in Section I, MOEAs produce a set of nondominated solutions in the final generation. All nondominated solutions have the same priority and cannot be compared. Nevertheless, it is necessary to select a single solution from the final non-dominated set.

In case of supervised classification, identification of the final solution is a relatively easy because a labeled training set can channel this selection. Different methods are used for obtaining final solution. In [67], a validation dataset is used for measuring the performance of each nondominated solution on independent data. The final solution is selected on the basis of performance on the validation set. In [68], a combination of the objective functions, feature correlation, and feature versus class correlation, called relative overall correlation, is used to select the final feature subset from the nondominated front.

In unsupervised classification the selection of the final solution is more difficult because labeled data is not accessible. It is observed that in most of the references, authors have not explained the method in detail. Therefore, it is difficult to discuss its merits or possible limitations.

F. Performance measures

In the surveyed references it is found that various different performance measures have been used to evaluate the performance of underlying classification algorithm. The kappa index [1] provides an accurate description of the classifier performance. It can be believed that it is better than the classification ratio as it takes into account the per class error distribution. The other cost function estimates aspects such as the generalization capability or the classifier overfitting. In many references, 10-fold cross-validation analysis to the training patterns has been taken into account to define the second cost function [2]. The area under the curve (AUC) of the receiver operating characteristics (ROC) has been reported as a better measure than accuracy for evaluating learning algorithms [23]. It compares classifier performance across the entire range of class distributions and error costs. It is found statistically consistent and more discriminative than accuracy. In [37], Jim´enez, F et al. compared ENORA and NSGA-II in terms of hypervolume statistics of the last population. This selection algorithm returns the best from two random individuals according to a rank-crowding-better function, by means of which an individual I is considered better than an individual J if its rank is better (lower) than the rank of the individual J in the population P.

G. Relative Comparison and Applications

To facilitate the reading all reviewed applications based on feature selection using MOEAs are enlisted in Table I. Tabsle II presents comparison of different MOEAs for Feature Selection.

V. DISCUSSION

It is observed that most of the Multiobjective Evolutionary Algorithms reported in the references depend on the standard crossover and mutation operators. They rarely use refined or hybrid operators. These refined operators may improve the quality of offspring which further lead to improvement in classification accuracy. The comparisons between standard and refined operators have not done in any work.

Multiobjective Evolutionary Algorithms generate a set of nondominated solutions in the final generation. It is essential to select a single solution from the final nondominated set. Although a number of different approaches have been proposed for selecting the final solution from the nondominated front, none of the solution selection methods have tried to combine this information with all the nondominated solutions through some kind of ensemble.

The problem of premature convergence is one of the most important parameters in case of evolutionary algorithms. To overcome this problem and to maintain diversity in the population some explicit measures has to be addressed. To the best knowledge, such kinds of actions are not reported in the available literature.

The results of feature selection depend on the selected classification algorithm. Additionally, the number of objective functions and their choice play an important role in the selection of the final feature subset. Therefore it becomes important to perform a comparative study of the performance of the proposed techniques based on some benchmark data sets. To the best knowledge, no comparative study of this manner has been reported so far in the literature.

VI. CONCLUSION

Most data mining and machine learning tasks aim to optimize multiple model parameters simultaneously. To deal with such tasks Multi-objective Evolutionary Algorithms become a popular choice. Since 2007 with the increase in the volume of data generation in data mining and machine learning applications MOEAs become very well-liked by researchers. In the past decade variety of data mining applications adopted MOEAs but, this review observes most of the fundamental aspects are yet need to be handled. No work has reviewed the papers which have been published in recent five years. Therefore this work mostly reviews the papers published since 2011 to 2017. Different aspects of MOEAs for feature selection such as adopted MOEAs, chromosome representation, objective functions, evolutionary operators, selecting the final solution, performance measures are described in this work. Also, a detailed review containing year, MOEA, new algorithms, crossover, mutation, future scope and limitation of each algorithm is provided in chronological order.

TABLE I. APPLICATIONS BASED ON MOEAS FOR FEATURE SELECTION

Sr. No.	Applications	References
1.	EEG Classification for Brain Computer Interfaces (BCI)	[1]
2.	Network Anomaly Detection	[3]
3.	Intrsion Detection System	[23]
4.	Bankruptcy Prediction	[43]
5.	Online Sales Forecasting	[62]
6.	Handwritten Digit Recognition	[52]
7.	Cervix Lesion Classification	[61]
8.	Behaviour Assessment of Children	[60]
9.	Cancer Classification	[47]
10.	Handwritten Month Word Recognition	[38]
11.	Churn Prediction in Telecommunication Service	[42]
12.	Multi-Machine Flow Shop Scheduling Problem Under Blocking	[63]
13.	Cardiac SPECT Diagnosis	[40]
14.	Multi-Skill Contact Center	[59]
15.	Human Motion Detection & Classification	[53]
16.	Gene Expression	[55]
17.	Sensor Event Mining	[57]
18.	Named Entity Recognition	[45]
19.	Anaphora Resolution	[46]

VII. REFERENCES

[1] Martín-Smith, Pedro, et al. "A Label-Aided Filter Method for Multi-objective Feature Selection in EEG Classification for BCI." *International Work-Conference on Artificial Neural Networks*. Springer International Publishing, 2015. DOI: 10.1007/978-3-319-19258-1_12

[2] Das, Ayan, and Swagatam Das. "Feature Weighting and Selection with a Pareto-optimal Trade-off between Relevancy and Redundancy." *Pattern Recognition Letters* (2017). http://dx.doi.org/10.1016/j.patrec.2017.01.004

[3] De la Hoz, Emiro, et al. "Feature selection by multi-objective optimisation: Application to network anomaly detection by hierarchical self-organising maps." *Knowledge-Based Systems* 71 (2014): 322-338. http://dx.doi.org/10.1016/j.knosys.2014.08.013

[4] Mukhopadhyay, Anirban, et al. "A survey of multiobjective evolutionary algorithms for data mining: Part I." IEEE Transactions on Evolutionary Computation 18.1 (2014): 4-19.

[5] B. Xue, M. Zhang, W. N. Browne and X. Yao, "A Survey on Evolutionary Computation Approaches to Feature Selection," in *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 4, pp. 606-626, Aug.2016. doi: 10.1109/TEVC.2015.2504420

[6] Ahuja, Jyoti, and Saroj Dahiya Ratnoo. "Feature Selection using Multi-objective Genetic Algorith m: A Hybrid Approach." INFOCOMP Journal of Computer Science 14.1 (2015): 26-37.

[7] U. Maulik, S. Bandyopadhyay, and A. Mukhopadhyay, Multiobjective Genetic Algorithms for Clustering—Applications in Data Mining and Bioinformatics. Berlin, Germany: Springer, 2011.

[8] K. Deb, Multi-Objective Optimization Using Evolutionary Algorithms London, U.K.: Wiley, 2001.

[9] C. A. Coello Coello, G. B. Lamont, and D. A. van Veldhuizen, Evolutionary Algorithms for Solving Multi-Objective Problems (Genetic and Evolutionary Computation), 2nd ed. Berlin/Heidelberg, Germany: Springer, 2007

[10] Pappa, Gisele L., Alex A. Freitas, and Celso AA Kaestner. "Attribute selection with a multi-objective genetic algorithm." Brazilian Symposium on Artificial Intelligence. Springer Berlin Heidelberg, 2002.

[11] C. M. Fonseca and P. J. Fleming, "Genetic algorithms for multiobjective optimization: Formulation, discussion and generalization," in Proc. 5th Int. Conf. Genet. Algorithms, 1993, pp. 416–423.

[12] J. Horn and N. Nafpliotis, "Multiobjective optimization using the niched Pareto genetic algorithm," Univ. Illinois

at Urbana-Champaign, Urbana, IL, USA, Tech. Rep. IlliGAl Rep. 93005, 1993.

[13] N. Srinivas and K. Deb, "Multiobjective optimization using nondominated sorting in genetic algorithms," Evol. Comput., vol. 2, no. 3, pp. 221–248, 1994.

[14] E. Zitzler and L. Thiele, "Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach," IEEE Trans. Evol. Comput., vol. 3, no. 4, pp. 257–271, Nov. 1999.

[15] E.Zitzler,M.Laumanns,andL.Thiele,"SPEA2:Improvingth estrength Pareto evolutionary algorithm," in Proc. EUROGEN, 2001, pp. 95–100.

[16] J. D. Knowlesand, D. W. Corne,"The Pareto archived evolution strategy: A new baseline algorithm for Pareto multiobjective optimisation," in Proc. IEEE Cong. Evol. Comput., 1999, pp. 98–105.

[17] D. W. Corne, J. D. Knowles, and M. J. Oates, "The Pareto envelope based selection algorithm for multiobjective optimization," in Proc. Conf. PPSN-VI, 2000, pp. 839–848.

[18] D. W. Corne, N. R. Jerram, J. D. Knowles, and M. J. Oates, "PESA-II: Region-based selection in evolutionary multiobjective optimization," in Proc. GECCO, 2001, pp. 283–290.

[19] K. Deb, A. Pratap, S. Agrawal, and T. Meyarivan, "A fast and elitist multi objective genetic algorithm: NSGA-II,"IEEETrans.Evol.Comput. vol. 6, no. 2, pp. 182–197, Apr. 2002.

[20] E. Zitzler and S. K¨unzli, "Indicator-based selection in multiobjective search," in Proc. PPSN VIII, vol. 3242. Sep. 2004, pp. 832–842.

[21] M. Dash and H. Liu, Feature selection for classification.Intelligent Data Analysis, 1997. 1(1-4): p. 131-156.

[22] Wang, Zhichun, Minqiang Li, and Juanzi Li. "A multi-objective evolutionary algorithm for feature selection based on mutual information with a new redundancy measure." *Information Sciences* 307 (2015): 73-88. http://dx.doi.org/10.1016/j.ins.2015.02.031

[23] Zhu, Yingying, et al. "An improved NSGA-III algorithm for feature selection used in intrusion detection." *Knowledge-Based Systems* 116 (2017): 74-85. http://dx.doi.org/10.1016/j.knosys.2016.10.030

[24] S. Theodoridis, K. Koutroumbas, Pattern Recognition, Academic Press 2009.

[25] W. Navidi, Statistics for Engineers and Scientists, third edition, McGraw-Hill, 2010.

[26] J. Quinlan, Induction of decision trees, Machine Learning 1 (1) (1986) 81–106.

[27] M.A. Hall, Correlation-based feature selection for discrete and numeric class machine learning, in: Proceedings of the Seventeenth International Conference on Machine Learning, ICML'00, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (2000) 359–366.

[28] L. Yu, H. Liu, Efficient feature selection via analysis of relevance and redundancy, Journal of Machine Learning Research 5 (2004) 1205–1224.

[29] M. Dash and H. Liu, "Feature selection for classification,"Intell. Data Anal., vol. 1, nos. 1–4, pp. 131–156, 1997

[30] Y. Liu, F. Tang, and Z. Zeng, "Feature selection based on dependency margin,"IEEE Trans. Cybern., vol. 45, no. 6, pp. 1209–1221, Jun. 2015.

[31] H. Liu and Z. Zhao, "Manipulating data and dimension reduction methods: Feature selection," in Encyclopedia of Complexity and Systems Science. Berlin, Germany: Springer, 2009, pp. 5348–5359.

[32] H. Liu, H. Motoda, R. Setiono, and Z. Zhao, "Feature selection: An ever evolving frontier in data mining," inProc. JMLR Feature Sel. Data Min., vol. 10. Hyderabad, India, 2010, pp. 4–13.

[33] H. Liu and L. Yu, "Toward integrating feature selection algorithms for classification and clustering", IEEE Trans. Knowl. Data Eng., vol. 17, no. 4, pp. 491–502, Apr. 2005.

[34] Yong, Zhang, Gong Dun-wei, and Zhang Wan-qiu. "Feature selection of unreliable data using an improved multi-objective PSO algorithm." *Neurocomputing* 171 (2016): 1281-1290. doi>10.1016/j.neucom.2015.07.057

[35] Ahuja, Jyoti, and Saroj Dahiya Ratnoo. "Feature Selection using Multi-objective Genetic Algorith m: A Hybrid Approach." INFOCOMP Journal of Computer Science 14.1 (2015): 26-37.

[36] Emmanouilidis, Christos, et al., "A multi-objective genetic algorithm approach to feature selection in neural and fuzzy modeling." Evolutionary Optimization3.1 (2001): 1-26.

[37] Pappa, Gisele L., Alex A. Freitas, and Celso AA Kaestner. "Attribute selection with a multi-objective genetic algorithm." Brazilian Symposium on Artificial Intelligence. Springer Berlin Heidelberg, 2002.

[38] Oliveira, Luiz S., Marisa Morita, and Robert Sabourin. "Feature selection for ensembles using the multi-objective optimization approach." *Multi-Objective Machine Learning*. Springer Berlin Heidelberg, 2006. 49-74.

[39] D. Zaharie, D. Lungeanu, S. Holban, "Feature ranking based on weights estimated by multiobjective optimization", *Proceedings of IADIS First European Conference on Data Mining*, pp. 124-128, 2007.

[40] Gaspar-Cunha, António. "Feature selection using multi-objective evolutionary algorithms: application to cardiac SPECT diagnosis." *Advances in Bioinformatics*. Springer

Berlin Heidelberg, 2010. 85-92. 10.1007/978-3-642-13214-8_11

[41] Waqas, Kashif, Rauf Baig, and Shahid Ali. "Feature subset selection using multi-objective genetic algorithms." *Multitopic Conference, 2009. INMIC 2009. IEEE 13th International*. IEEE, 2009. DOI: 10.1109/INMIC.2009.5383159

[42] Huang, Bingquan, Brian Buckley, and T-M. Kechadi. "Multi-objective feature selection by using NSGA-II for customer churn prediction in telecommunications." *Expert Systems with Applications* 37.5 (2010): 3638-3646.

[43] Gaspar-Cunha, António, et al. "Multi-objective evolutionary algorithms for feature selection: application in bankruptcy prediction." *Asia-Pacific Conference on Simulated Evolution and Learning*. Springer Berlin Heidelberg, 2010. DOI: 10.1007/978-3-642-17298-4_3

[44] Venkatadri, M., and K. Srinivasa Rao. "A multiobjective genetic algorithm for feature selection in data mining." International Journal of Computer Science and Information Technologies 1.5 (2010): 443-448.

[45] Ekbal, Asif, and Sriparna Saha. "Multiobjective optimization for classifier ensemble and feature selection: an application to named entity recognition." *International journal on document analysis and recognition* 15.2 (2012): 143-166. doi:10.1007/s10032-011-0155-7

[46] Ekbal, Asif, et al. "Multiobjective simulated annealing based approach for feature selection in anaphora resolution." *Discourse Anaphora and Anaphor Resolution Colloquium*. Springer Berlin Heidelberg, 2011. DOI: 10.1007/978-3-642-25917-3_5

[47] Dash, Sujata, Bichitrananda Patra, and B. K. Tripathy. "Study of Classification Accuracy of Microarray Data for Cancer Classification using Multivariate and Hybrid Feature Selection Method." IOSR Journal of Engineering (IOSRJEN) 2.8 (2012): 112-119.

[48] Li, Xiangtao, and Minghao Yin. "Multiobjective binary biogeography based optimization for feature selection using gene expression data." *IEEE Transactions on NanoBioscience* 12.4 (2013): 343-353. DOI: 10.1109/TNB.2013.2294716

[49] Xue, Bing, Mengjie Zhang, and Will N. Browne. "Particle swarm optimization for feature selection in classification: A multi-objective approach." *IEEE transactions on cybernetics* 43.6 (2013): 1656-1671.

[50] Paul, Sujoy, and Swagatam Das. "Simultaneous feature selection and weighting–An evolutionary multi-objective optimization approach." *Pattern Recognition Letters* 65 (2015): 51-59.

[51] Xue, Bing, Wenlong Fu, and Mengjie Zhang. "Multi-objective Feature Selection in Classification: A Differential Evolution Approach." *SEAL*. 2014. DOI:10.1007/978-3-319-13563-2_44

[52] Oliveira, Luiz S., et al. "Multi-objective genetic algorithms to create ensemble of classifiers." *International Conference on Evolutionary Multi-Criterion Optimization*. Springer Berlin Heidelberg, 2005. DOI: 10.1007/978-3-540-31880-4_41

[53] Tan, Choo Jun, Chee Peng Lim, and Yu–N. Cheah, "A multi-objective evolutionary algorithm-based ensemble optimizer for feature selection and classification with neural network models." *Neurocomputing* 125 (2014): 217-228.

[54] Panda, Mrutyunjaya. "Big Models for Big Data using Multi objective averaged one dependence estimators." arXiv preprint arXiv:1610.07752(2016).

[55] Smith, Christopher, and Yaochu Jin. "Evolutionary multi-objective generation of recurrent neural network ensembles for time series prediction." Neurocomputing 143 (2014): 302-311.

[56] Anusha, M., and J. G. R. Sathiaseelan. "Feature selection using k-means genetic algorithm for multi-objective optimization." Procedia Computer Science 57 (2015): 1074-1080.

[57] Mehdiyev, Nijat, et al. "Sensor event mining with hybrid ensemble learning and evolutionary feature subset selection model." *Big Data (Big Data), 2015 IEEE International Conference on*. IEEE, 2015.

[58] Khan, Ayesha, and Abdul Rauf Baig. "Multi-Objective Feature Subset Selection using Non-dominated Sorting Genetic Algorithm." *Journal of applied research and technology* 13.1 (2015): 145-159. http://dx.doi.org/10.1016/S1665-6423(15)30013-4

[59] Jiménez, Fernando, et al. "Attribute Selection Via Multi-Objective Evolutionary Computation Applied to Multi-Skill Contact Center Data Classification." Computational Intelligence, 2015 IEEE Symposium Series on. IEEE, 2015

[60] Jiménez, F., et al. "Multi-Objective Evolutionary Computation Based Feature Selection Applied to Behaviour Assessment of Children." *World Academy of Science, Engineering and Technology, International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering* 10.6 (2016): 2065-2073.

[61] Sahoo, Anita, and Satish Chandra. "Multi-Objective Grey Wolf Optimizer for Improved Cervix Lesion Classification." *Applied Soft Computing* (2016). http://dx.doi.org/10.1016/j.asoc.2016.12.022

[62] Jiménez, F., et al. "Multi-objective evolutionary feature selection for online sales forecasting." *Neurocomputing* (2016). http://dx.doi.org/10.1016/j.neucom.2016.12.045

[63] Rosales-Pérez, Alejandroa; * | Gonzalez, Jesus A.b | Coello, Carlos A. Coelloc | Reyes-Garcia, Carlos A.b | Escalante, Hugo Jairb Intelligent Data Analysis, vol. 20, no. s1, pp. S37-S51, 2016 10.3233/IDA-16084

[64] Nouri, Nouha, and Talel Ladhari. "Evolutionary multiobjective optimization for the multi-machine flow shop scheduling problem under blocking." *Annals of Operations Research* (2017): 1-18

[65] Cai, Fuyu, et al. "Fuzzy Criteria in Multi-objective Feature Selection for Unsupervised Learning." *Procedia Computer Science* 102 (2016): 51-58. https://doi.org/10.1016/j.procs.2016.09.369

[66] Xue, Bing, et al. "Multi-objective evolutionary algorithms for filter based feature selection in classification." *International Journal on Artificial Intelligence Tools* 22.04 (2013): 1350024. DOI: 10.1142/S0218213013500243

[67] L. E. S. de Oliveira, R. Sabourin, F. Bortolozzi, and C. Y. Suen, "A methodology for feature selection using multiobjective genetic algorithms for handwritten digit string recognition,"Int. J. Pattern Recognit. Artif. Intell., vol. 17, no. 6, pp. 903–929, 2003

[68] C.-M. Wang and Y.-F. Huang, "Evolutionary-based feature selection approaches with new criteria for data mining: A case study of credit approval data," Expert Syst. Appl., vol. 36, no. 3, pp. 5900–5908, Apr. 2009.

TABLE II. COMPARISON OF DIFFERENT MOEAS FOR FEATURE SELECTION

Author/ Year	Algorithm	MOEA used	Classification	Objective Functions	Type (W/F/E)/ Measure	Evolutionary Operators	Future Scope/Limitations
Christos Emmanouilidis et. al. [36] 2001	MOGA in Fuzzy Neural Modelling	MOGA	Supervised	number of features,classifier accuracy	wrapper, MLP	one point crossover, bit flip mutatuion	apply objectives such as true positive, true negative classification rates, data acquisition or misclassification costs.
Gisele L. Pappa et. al. [37] 2002	MOGA for Attribute Selection	MOGA	Supervised	classifier accuracy, computational complexity	wrapper, C4.5	one point crossover, bit flip mutatuion	use of a niching method to achieve greater population diversity for reducing premature convergence
Luiz Oliveira et. al. [38] 2006	FS for ensembles	hierarchical MOGA	supervised, unsupervised	number of features,classifier accuracy	wrapper MLP, HMM	one point crossover, bit flip mutatuion	-
Daniela Zaharie et. al. [39] 2007	Feature Weighing based MOEA	NSGA-II	supervised, unsupervised	number of features,classifier accuracy	filter, intra-class dissimilarity, inter class dissimilarity, attribute-class correlation	one point crossover, bit flip mutatuion	integration of this pre-processing technique into a system for risk prediction in obstetrics.
António Gaspar Cunha et. al. [40] 2007	RPSGAe + SVM	RPSGAe	supervised	number of features,classifier accuracy	wrapper, SVM	one point crossover, bit flip mutatuion	-
Kashif Waqas et. al. [41] 2009	FSS-MOGA	NSGA	supervised	classifier accuracy, generalization capability	-	one point crossover, bit flip mutatuion	usage of Ensembles, Preference ordering ranking genetic algorithm (POGA)
Bingquan Huang et. al. [42] 2009	FBSM	NSGA-II	supervised	overall accuracy,accuracy of true churn and true nonchurn	wrapper, Decision Tree	one point crossover, bit flip mutatuion	to reduce computational overhead of the present approach and sampling technique of dataset
Ant´onio Gaspar-Cunha et. al. [43] 2010	RPSGA Feature Selection	RPSGA	supervised	classifier accuracy, classifier parameters	wrapper, SVM	one point crossover, bit flip mutatuion	-
Venkatadri.M et. al. [44] 2010	MOGA with	MOGA	supervised	number of features,classifier accuracy	filter, measures consistency dependency, distance, information	one point crossover, bit flip mutatuion	comparison of the proposed MOGA with other techniques

Asif Ekbal et. al. [45] 2011	MOO based ensamble	NSGA-II	supervised	number of features,classifier accuracy	wrapper, SVM, entropy	one point crossover, bit flip mutatuion	development of some vote-based classifier ensembles
Asif Ekbal et. al. [46] 2011	GA+ Simulated Annealing	MOGA	supervised	number of features,classifier accuracy	Simulated Annealing	one point crossover, bit flip mutatuion	comparison to the previously developed multiobjective genetic algorithm based feature selection technique.
Sujata Dash et. al. [47] 2012	CFS	-	supervised	number of features,classifier accuracy	filter: degree of redundenc, wrapper: J48, Random Forest and Random Trees	-	using algorithm in applications
Xiangtao Li et. al. [48] 2013	Hybrid MOBBBO + SVM	MOBBBO	supervised	number of features,classifier accuracy	wrapper, SVM	one point crossover, probabilistic mutatuion	algorithm applied to some problems in other fields
Bing Xue et. al. [63] 2013	NSPSOFS CMDPSOFS	Pareto Optimality	supervised	number of features,classifier accuracy	filter, hyper volume indicator	one point crossover, crowding mutatuion	investigate whether given can be used in wrapper method
Sujoy Paul et. al. [64] 2013	MOEA/D	MOEA	supervised	inter-class intraclass distances	wrapper, k-NN	one point crossover, bit flip mutatuion	development of a classifier, which exploits the intra- and inter-class distance property of the selected and weighted subset of features
Emiro de la Hoz et. al. [3] 2014	GHSOMs	NSGA-II	unsupervised	classifier accuracy, generalization capability	filter, Jaccard's coefficient	one point crossover, bit flip mutatuion	hybridization of Gaussian Mixture Model and Support Vector Machines
Bing Xue et. al. [51] 2014	DEFS, DEFS2, DEMOFS	NSGA-II	supervised	number of features,classificat ion error rate	wrapper, k-nearest neighbour	one point crossover, bit flip mutatuion	binary Differential Evolution algorithm, using filter approach
Luiz S. Oliveira et. al. [52] 2014	MOGA based FS	MOGA	supervised, unsupervised	number of features, number of classifiers, accuracy	wrapper, multi-layer perceptron (MLP) neural networks	one point crossover, bit flip mutatuion	make the issue of using diversity to build ensembles
Choo Jun Tan et. al. [53] 2014	MmGA-based ensemble optimizer	MmGA	supervised	number of features,classifier accuracy	wrapper, MLP, RBF	one point crossover, bit flip mutatuion	inverted generational distance and spread will be used for formulating the elite-selection scheme and for tracking the behaviour of the optimizers
Mrutyunjaya Panda et. al. [54] 2014	AODE with ENORA	ENORA	supervised	number of features,classifier accuracy	wrapper, AODE	one point crossover, bit flip mutatuion	deep broad learning for big data with new nature inspired algorithms
Christopher Smith et. al. [55] 2014	Hybrid-MOEA for RNN	Hybrid-MOEA	supervised	complexity, training errors	wrapper, RNN	one point crossover, bit flip mutatuion	HMOEA and ensemble member selection methods to complex engineering problems
M.Anusha et. al. [56] 2014	Extended NLMOGA	NLMOGA	unsupervised	diversity, comp actness	wrapper, k-means	one point crossover, probability mutatuion	improve algorithm with constrained crossover on high dimensional data sets
Nijat Mehdiyev et. al. [57] 2014	Hybrid Ensemble Learning using ENORA	ENORA	supervised	number of features,classifier accuracy	filter	one point crossover, bit flip mutatuion	algorithms will be extended with other applications
P Martín-Smith et. al. [1] 2015	Label-Aided Filter Approach	NSGA-II	supervised, unsupervised	classifier accuracy, generalization	filter, coincidence measure	simulated binary crossover,	analysis of characteristics of the featuresselected for obtaining knowledge about important

				capability		mutation	electrodes and segments
Zhichun Wang et. al. [22] 2015	MECY-FS	SPEA2	supervised	feature relevance and redundancy	hybrid	variant of uniform crossover, mutation	using wrapper approach
Zhang Yong et. al. [34] 2015	BMOPSOFS	PSO with crowding distance	supervised	reliability, classification accuracy	meta-heuristic	hybrid mutatuion	-
Jyoti Ahuja et. al. [35] 2015	MOGA: A hybrid approach	MOGA at filter, GA at wrapper	supervised	intercorrelation, intracorrelation	hybrid	one point crossover, bit flip mutatuion	reserachers can select preferences in feature selection at the end of the filter stage
A. Khan et. al. [58] 2015	FSS using NSGA-II	NSGA-II	supervised	number of features,classifier accuracy	wrapper, ID3	one point crossover, bit flip mutatuion	using other multiobjective algorithms
Fernando Jim´enez et. al. [59] 2015	ENORA to Multi-Skill Contact Center	ENORA	supervised	number of features,classifier accuracy	wrapper, C4.5	Adaptive crossover, Adaptive mutatuion	-
F. Jim´enez et. al. [60] 2016	BASC-II using MOEA	ENORA, NSGA-II	unsupervised	likelihood of cluster,classifier accuracy	wrapper, C4.5	one point crossover, bit flip mutatuion	-
Anita Sahoo et. al. [61] 2016	MOGWO, NSGWO	MOGA NSGA-II	supervised	dimensionality of feature subset,classifier accuracy	wrapper, SVM	one point crossover, bit flip mutatuion	identify the most suitable meta-heuristic in Non-dominated Sorting based framework
F. Jim´enez et. al. [62] 2016	ENORA based FS	ENORA	supervised	number of features, root mean squared error	wrapper, Random Forest	one point crossover, bit flip mutatuion	incorporation of ENORA as search strategy in multivariate filters and in other heuristic search algorithms like PSO
Alejandro Rosales-P´erez et. al. [63]	EMOPG + FS	PAES	supervised	number of features,classifier accuracy	wrapper, k-NN	one point crossover, bit flip mutatuion	testing on large scale data sets
Nouha Nouri et. al. [64] 2016	MOEA for Flow Shop Scheduling	MOGA	-	makespan, total completion time	-	two point crossover, shift mutatuion	conduct a comparative study between existing multiobjective evolutionary algorithms
Fuyu Cai et. al. [65] 2016	FC-MOFS	NSGA-II	unsupervised	number of features,classifier accuracy	hybrid	one point crossover, bit flip mutatuion	comprehensive and systematic validation considering different combinations of clustering algorithms and objective functions
Ayas Das et. al. [2] 2017	Feature selection & weighing	MOEA/D	supervised	relevancy, redundancy	wrapper	one point crossover, bit flip mutatuion	improvement on the sparsity penalty
Yingying Zhu et. al. [4] 2017	I-NSGA-III	NSGA-III	supervised	classifier accuracy, computational complexity	wrapper, Decision Tree	one point crossover, bit flip mutatuion	-
Bing Xue et. al. [66] 2017	NSGAIIMI NSGAIIE SPEA2MI SPEA2E	NSGA-II, SPEA 2	supervised	classifier accuracy, number of features	filter	one point crossover, bit flip mutatuion	application on proposed algorithms

IJEAST

INTERNATIONAL JOURNAL OF ENGINEERING APPLIED SCIENCE AND TECHNOLOGY

ABOUT IJEAST

International Journal of Engineering Applied Science and Technology (IJEAST) is a peer-reviewed, open access journal that publishes high-quality research papers in the field of Engineering, Applied Science and Technology.

IJEAST aims to provide a platform for researchers, academicians, and professionals to share their innovative ideas, research findings, and practical experiences with the global scientific community.

FOCUS AREAS

- Engineering
- Applied Science
- Technology
- Innovation & Development
- Interdisciplinary Studies

PEER REVIEWED

All submissions are rigorously peer reviewed to ensure quality.

OPEN ACCESS

Free and unrestricted access to research for all.

GLOBAL REACH

Connecting researchers and professionals worldwide.

TIMELY PUBLICATION

We ensure a swift and efficient publication process.

For more information, visit our website
www.ijeast.com

IJEAST

INTERNATIONAL JOURNAL
OF ENGINEERING APPLIED SCIENCE
AND TECHNOLOGY

editor@ijeast.com
www.ijeast.com
India

2455-2143