Multi-objective Evolutionary Algorithms for Classification: A Review

Seema Mane¹, Prof. S. S. Sonawani² Dr. Sachin Sakhare³ and Prof. P. V. Kulkarni⁴

¹Master in Computer Engineering, Department of Computer Engineering, MIT, Pune, Pune University
Gadiya Estate, S/n: 81/6, Paud Road, Kothrud Depo, Pune 038

²Assistant Professor, Department of Computer Engineering, MIT, Pune, Pune University

³Head of Department, Department of Computer Engineering, VIT, Pune, Pune University

⁴Assistant Professor, Department of Computer Engineering, MIT, Pune, Pune University

ABSTRACT
Multi-objective evolutionary algorithms are evolutionary systems which are used for optimizing various measures of the evolving systems. Most of the real life data mining problems are optimization problems, where the aim is to evolve a candidate model that optimizes certain performance criteria. Classification problem can be thought of as multi-objective problem as it may require to optimize accuracy, model complexity, interestingness, misclassification rate, sensitivity, specificity etc. The performance of these MOEAs used is depends on various characteristics like evolutionary techniques used, chromosome representation, parameters like population size, crossover rate, mutation rate, stopping criteria, number of generations, objectives taken for optimization, fitness function used, optimization strategy etc. This paper reports the comprehensive survey on recent developments in the multi-objective evolutionary algorithms for classification problems.

Keywords:- Multi-objective Optimization, Evolutionary algorithms, Classification, Pareto optimality, Genetic Algorithm.

1. INTRODUCTION
Data Mining is the process of discovering interesting and important patterns from large datasets. The main objective of any data mining process is to build an efficient predictive or descriptive model of huge volume of data. This model must best fit the data also able to generalize to new data. There are different types of tasks associated to data mining process i.e. classification, clustering, association rule mining etc. Due to the complexity of classification problems, cannot be solved using standard mathematical techniques. Evolutionary algorithms have been found to be useful in automatic processing of large volume of raw noisy data due to their inherent parallel architecture [1], [2]. Traditionally, evolutionary algorithms were used to solve the single-objective classification problems. But there are many real life classification problems having multiple conflicting objectives, which need to be optimized simultaneously to obtain optimal solutions. Conflicting objectives for classification problems can be accuracy, sensitivity, specificity, misclassification rate, mean squared error etc. Therefore, the concept of multi-objective optimization is highly applicable to classification problems. Classification problems are multi-objective in nature, and the goal is to simultaneously optimize all conflicting objectives. Optimum performance in one objective results in low performance in one or more of the other objectives, creating necessity for compromise [1]. For solving multi-objective optimization problems, number of evolutionary algorithms (EAs) have been proposed in the literature [ref should add]. In single-objective optimization single optimum solution is generated in final generation, while in multi-objective optimization, set of non-dominated solutions are generated in final generation, where each objective can be improved only by degradation of at least one of the other objectives. Multi-objective evolutionary algorithms (MOEAs) [3] have become popular in data mining. MOEAs can be found in the literature for solving data mining task such as classification, clustering, association rule mining. A variety of MOEAs used for different classification problems like in medical field. Different classification techniques such as Decision tree, Rule based classifier, neural network, Bayesian network, Support vector machine can be used for predicting class labels for unknown tuples. In this paper, we attempt to make comprehensive survey of the recent developments in the MOEAs for solving different classification problems. Section 2 gives detailed introduction of multi-objective optimization. Evolutionary computing approaches for multi-objective optimization provided in section 3. Issues need to be considered during implementation of multi-objective evolutionary algorithms are given in section 4. Section 5 gives review of different classification problems solved using different multi-objective optimization. Conclusion and Future scope drawn in section 6.
2. MULTI-OBJECTIVE OPTIMIZATION

This section introduced some basic concepts of MOO. Then, an overview of available MOEAs is provided.

2.1 Basic Concepts of Multi-objective Optimization

In real life situations, there may be multiple objectives need to be optimized simultaneously in order to solve certain classification problems. Classification problems are multi-objective in nature as they required simultaneously optimization of multiple objectives like accuracy, sensitivity, mean squared error etc. The main difficulty with multi-objective optimization is that there is no accepted predefined definition of optimum, so it is difficult to compare one solution with another one. Multi-Objective Optimization Problem (MOOP), in general may be stated as, finding the value for a set of n decision variables which must satisfy some constraints (I inequalities and K equalities) such that the M objective functions are optimized and can be modelled as follows [3]:

Maximize/Minimize \( f_m(x), \) 
Subject to \( g_j(x) \leq 0, j=1,2,3,\ldots,J; \)
\( h_k(x) = 0, k=1,2,3,\ldots,K; \)
\( x_i^L \leq x_i \leq x_i^U, i=1,2,3,\ldots,n. \)

A solution \( x \) is vector of \( n \) decision variables: \( x = (x_1, x_2, x_3, \ldots, x_n)^T. \) The solutions which satisfy the constraints and variable bounds constitute a feasible decision variable space \( S. \) \( g_j \) is the set of inequality constraints and \( h_k \) is set of equality constraints. Multi-objective optimization objective functions constitute multi-dimensional space, in addition to the usual decision variable space. The additional space is called objective space \( Z. \) For each solution \( x \) in the decision variable space, there exists a point in the objective space, denoted by \( f(x) = z = (z_1, z_2, \ldots, z_M)^T. \) The mapping takes place between \( n \)-dimensional solution vector and \( M \)-dimensional objective vector.

We introduce some definitions to describe the concept of optimality.

Definition 1 (Dominance) [3]: A solution \( x^{(1)} \) is said to dominate the other solution \( x^{(2)}, x^{(1)} \leq x^{(2)}, \) if following both conditions are satisfied:

1. if \( x^{(1)} \) is no worse than \( x^{(2)} \) in all objectives and
2. \( x^{(1)} \) is strictly better than \( x^{(2)} \) in at least one objective.

If any of the above conditions violated, the solution \( x^{(1)} \) does not dominate solution \( x^{(2)}. \)

Definition 2 (Non-dominated set) [3]: Among the set of solutions \( P, \) the non-dominated solutions are those \( P' \) which are not dominated by any member of the set \( P. \)

Definition 3 (Globally Pareto-optimal set) [3]: The non-dominated set of the entire feasible search space \( S \) is the Globally Pareto-optimal set.

Definition 4 (Strong Dominance) [3]: A solution \( x^{(1)} \) strongly dominates solution \( x^{(2)}, \) if solution \( x^{(1)} \) is strictly better than solution \( x^{(2)} \) in all M objectives.

Definition 5 (Weak Dominance) [3]: Among the set of solutions \( P, \) weakly non-dominated set of solutions \( P' \) are those that are not strongly dominated by any member of set \( P. \)

2.2 Multi-objective Evolutionary Algorithms

Traditional search and optimization methods such as gradient-based methods are difficult to extend to the multi-objective concept because their basic design excludes the consideration of multiple solutions. In contrast, evolutionary algorithms are well-suited for handling such situations. There are different approaches for solving multi-objective optimization problems [3, 4]. MOEAs have evolved over several years, starting from conventional weighted formula approach to the elitist Pareto approach. In the weighted formula approaches, multiple objective functions are combined into single scalar value using weights, and that single-objective function is then optimized using conventional evolutionary algorithms. Schaffer in 1984 implemented first multi-objective genetic algorithm to find the set of non-dominated solutions. In this population based non-elitist, non-pareto approach such as Vector Evaluated Genetic Algorithm (VEGA) [5], a selection operator is used and numbers of subpopulations are generated equal to the number of objectives to be optimized. Each individual of subpopulation assigned fitness with respect to the respective objective function. In Pareto-based approach, selection should be made using a non-dominated ranking scheme and that diversity should be maintained with the use of a sharing function. The most representative non-elitist Pareto-based MOEAs are multiple objective genetic algorithm (MOGA) [6], niched Pareto Genetic algorithm (NPGA) [7], non-dominated sorting genetic algorithm (NSGA) [8]. In MOGA [6], a ranking scheme is used to rank each individual corresponding to the number of individuals in the current population by which it is dominated. Fitness sharing is used in order to maintain diversity, with a mating restriction scheme to avoid crossover between very distant individuals in the search space. In NSGA [8], population is sorted in various fronts. Non-dominated individuals belonging to the front are more fit, hence they are removed from the population and the process is repeated until the entire population is classified. A tournament selection scheme based on Pareto dominance is used in NPGA [7]. These techniques do not use elitism, and therefore, they cannot give guarantee of preserving non-dominated solutions during search. In recent years, a number of elitist Pareto-based multi-objective evolutionary algorithms have been proposed. The most representative elitist MOEAs include Strength Pareto evolutionary algorithm (SPEA) [9], SPEA2 [10], Pareto archived evolution strategy (PAES) [11], Pareto Envelope-based selection algorithm (PESA) [12], and PESA II [13], and non-
dominated sorting genetic algorithm-II (NSGA II) [14]. In SPEA [9], it introduces elitism by maintaining external population. This population maintains non-dominated solutions that are found from beginning of the simulation. At each generation newly found non-dominated solutions are compared with external population and resulting non-dominated solutions are preserved. SPEA2 [10] incorporate a fine-grained fitness assignment strategy, a density estimation technique and an enhanced archive truncation method. In PAES [11], an archive of non-dominated solutions is considered for maintaining population diversity. A newly generated offspring is compared with the archive to verify if it dominates any member of the archive. If yes, then the offspring enters the archive and is accepted as a new parent. In NSGA II [14], uses crowding distance for density estimation for each individual. Crowded distance of a solution is the average side-length of the cube enclosing the point without including any other point in the population. Solutions of the last accepted front are ranked according to the crowded distance. Each solution is assigned fitness equal to its Non-domination level. Binary tournament selection, recombination and mutation are used to create an offspring population. Classification is the one of the important task in data mining. It is a supervised technique which is used to predict class labels of unknown tuples. Different techniques which are used for classification are Decision tree classifier, Rule based classifier, Neural network, Bayesian network and Support vector machine. Most of the applications of MOEAs for classification problems have used one of these Pareto-based elitist approaches as their underlying optimization strategy.

3. EVOLUTIONARY COMPUTING FOR SOLVING MULTI-OBJECTIVE OPTIMIZATION PROBLEMS

A multi-objective optimization problem is that where multiple conflicting objectives must be optimized simultaneously. Evolutionary computing is branch of computer science, to solve problems it uses Darwin’s principle of evolution. It includes variety of algorithms like genetic algorithm, genetic programming, evolutionary algorithms, evolutionary programming, cultural algorithms, swarm intelligence etc. Evolutionary algorithms performs global search and are convenient for parallelization. This is robust search method and adapts to the environment and discovers interesting knowledge. Thus evolutionary algorithms are suitable for multi-objective optimization problem since they allow various objectives incorporated simultaneously in the solution. When we combine neural network with evolutionary algorithms it becomes evolutionary artificial neural network.

3.2 Genetic Algorithm

Genetic algorithm is one of the global search algorithms to find optimal solution to the problem that is based on principles like selection, crossover, and mutation. The biological simile for genetic algorithms is the evolution of new species by survival of the fittest individuals, as described by Charles Darwin. New individuals are generated by crossover of genetic information of two parents. A genetic algorithm tries to replicate the natural evolution process. Its purpose is to optimize a set of parameters. There are two basic operators of genetic algorithm to regenerate new individuals are mutation and crossover. Genetic algorithm starts with the random generation with the initial set of individuals i.e. Initial population. Population evolves for many generations, when algorithm finishes it returns best solution. Genetic algorithms are used particularly to the problems where it is extremely difficult or impossible to find exact solution, or some difficult problems to which exact solution may not be required. Genetic algorithm is used for classification rule mining, also it is used with neural network to obtain optimized neural network training model for classification. The output of genetic algorithm is set of “If. . . .Then.....” rules for rule mining and in case of neural network the output of genetic algorithm is optimized structure of neural network.

3.3 Genetic Algorithm for Classification Rule Mining

As genetic algorithm span large search space to find diverse set of solutions, it has been applied to optimization problems. The search space is nothing but encoded as chromosomes. The encoding style can be binary or value encoding. In binary encoding, data in search space is encoded as string of binary bits, while in value encoding data is represented as string of values. Genetic algorithm selects two parents for crossover, randomly a crossover point is chosen and the parts of the string starting at the crossover point are swap to produce new two individuals. Mutation is done just by negating a bit. In case of rule mining system the search space is dataset of records. The chromosome is a binary string of length equal to the product of the number of attributes and the values they take [15]. The output of a Genetic algorithm as a rule mining system for Michigan style approach is a simple “If...Then” rule for each individual, each rule representing a class. In the case of Pittsburg style approach the output is a complex “If...Then...Else If” rule which encodes the entire system of knowledge base [15].

3.4 Genetic Algorithm for Neural Network

Evolutionary algorithms like genetic algorithm can be used to train the neural networks, design functions of their neurons, choose their structure. There are some reasons why to use genetic algorithm to train neural network, one of the most important reason is that genetic algorithm train the network no matter how it is connected- whether it is feed-
back or feed-forward network. Algorithms like back-propagation only train certain restricted topologies and types of networks. On the other hand genetic algorithm can train any types of network also mixture of two types of networks.

4. **ISSUES IN EVOLUTIONARY MULTI-OBJECTIVE OPTIMIZATION**

4.1 **Chromosome Representation**

The overall performance of the algorithm immensely depends upon the chromosome representation. In genetic algorithm, chromosomes can be encoded using the binary encoding or value encoding techniques. Also systems use fixed length chromosome or variable length chromosome.

4.2 **Genetic Algorithm parameters**

The performance of the algorithm is influenced by different parameters like population size represent number of individuals in each generation, crossover probability which is probability of creating new individuals via selected crossover type of selected individuals, number of generations represent the number of iterations of whole algorithm, mutation probability which is probability of creating new individuals via mutation based on selected individual and stopping criteria which may be specified by the user it may be number of generations or if a satisfactory solution found.

4.3 **Objectives for Optimization**

Optimization objectives for classification can be like Accuracy, Specificity, Sensitivity, Misclassification Rate, Precision, mean squared error and Recall. Accuracy is the most commonly used objective for optimization.

4.4 **Reproduction Parameters**

Reproduction operators are Selection, Crossover, and Mutation. Selection means that two chromosome be selected from the population and crossover or mutation is applied. There are three types of selection mechanisms used in EMOO for classifications are Roulette wheel selection where each individual is given a chance to become a parent in proportion to its fitness evaluation; Tournament selection in which a group of parents is selected and a tournament is held to decide which of the individuals will be the parent; Fitness ranking where individuals are sorted in order of raw fitness and given ranks. Good ranked individuals are chosen as parents [15]. Crossover applied to the parents which are selected by choosing crossover points. There are three types of crossover are one-point crossover, two-point crossover and uniform crossover. Mutation change the gene of chromosomes by changing attribute value if value encoding is used or by negating bits is binary encoding is used.

4.5 **Data sets**

There are various datasets on which classification can be applied. Datasets contains the information collected from various disciplines like chemistry, biology, finance, medicine, pattern recognition, physics, space research, electricity problems, etc. Classification can be applied on any type of datasets which are available in different areas of application.

5. **EVOLUTIONARY MULTI-OBJECTIVE OPTIMIZATION SYSTEMS FOR CLASSIFICATION**

There are varieties of evolutionary algorithms used in solving various problems; most frequently used algorithm is genetic algorithm. Also there are varieties of MOEA’s which are used for solving multi-objective optimization problem. Matthew Butler and Ali Daniyal proposed multi-objective optimization for financial forecasting with an evolutionary artificial neural network [16]. Authors used neuro-evolutionary approach to make accurate predictions of the movement of stock market which is based on NEAT. In addition with NEAT approach greedy mutation operator used that deploys back-propagation based adaptation of the current model’s weight parameters [16]. Hussein A. Abbass proposed evolutionary artificial neural network approach for breast cancer diagnosis based on Pareto-differential evolution (PDE) [17] augmented with local search [18]. This is a multi-objective optimization problem with two objectives first is to minimize number of hidden units and second, to minimize error. The major advantage of evolutionary approach is that ability to escape local optima. The problem with back-propagation or other training algorithms is choice of a correct architecture [18]. This problem has been solved by using evolutionary approach. Back-propagation has been used as local search technique to overcome slow convergence of evolutionary approach. Author showed that this multi-objective approach for breast cancer diagnosis has better generalization and much lower computational cost [18]. S. Dehuri, S. Patnaik, A. Ghosh and R. Mall proposed elitist multi-objective genetic algorithm for classification rule generation [19]. Objectives to be considered for optimization were predictive accuracy, comprehensibility and interestingness of rules [19]. As these objectives conflict with each other, this makes it a multi-objective optimization problem. Single-objective optimization approach gives single solution this approach not suitable in case of multiple and conflicting objectives. In presence of multiple conflicting objective optimizations gives set of optimal solutions instead of single optimal solution. Authors proposed elitist MOGA [19] approach with hybrid crossover operator to extract classification rules of “If.......Then” form.
Multi-objective evolutionary approach applied for system identification with recurrent neural network by J.H.Ang, C.K.Goh, E.J.Tech and A.A.Mamun [20]. Authors incorporated few features such as variable-length chromosome representation in the form of structural mutation and micro genetic algorithm for local search in multi-objective evolutionary recurrent neural network. They considered simultaneous evolution of synaptic weights and neural network architecture. Objectives to be considered for optimization are accuracy and network complexity [20] which are conflicting with each other. Uniform mutation was used to evolve required set of connection weights. In this approach, elitism was implemented as fixed size archive to prevent loss of good individuals due to the stochastic nature of optimization process. This approach is effective for system identification, where hidden neurons were evolved using structural mutation and variable chromosomes was used to model network architecture. Renata Furtuna, Silvia Curteau and Florin Leon proposed an elitist non-dominated sorting genetic algorithm enhanced with a neural network applied to the multi-objective optimization of a polysiloxane synthesis process [21]. Objectives considered by authors for optimization are, first, is to maximize the reaction conversion and second is to minimize the difference between the obtained viscometric molecular weight and the desired molecular weight [21]. A feed-forward neural network was used with NSGA II [14]. NSGA II [14] approach was used for multi-objective optimization of polysiloxane synthesis process. This NSGA II approach use elitist mechanism to preserve good individuals for next generation as there may be chance of getting loss of good individuals due to optimization process. The real coding was used for chromosome encoding as it is more suited for real life problems. Binary tournament selection was used for selection of parents for new individual’s reproduction. Ranking method was used for selection of new population for next generation. Crowding distance was used to maintain diversity of solutions. This approach quickly gives optimal solutions as an acceptable compromise between objectives competes with each other [21]. Manuel Cruz-Ramírez, César Hervás-Martínez, Juan Carlos Fernández, Javier Briceño and Manuel de la Mata proposed multi-objective approach with evolutionary artificial neural networks for predicting patient survival after liver transplantation [22]. Objectives which are to be considered for optimization are Accuracy and minimum sensitivity. NSGA II [14], a multi-objective evolutionary algorithm was used to train radial basis function neural networks. The optimal neural network models obtained from Pareto fronts were used to give input for rule based system which was used to find perfect donor-recipient match. A major disadvantage of evolutionary approach is, it is computationally expensive and thus evolutionary approach is slow. Hybrid technique such as by augmenting evolutionary algorithm with local search was used to speed up slow convergence.

6. CONCLUSION AND FUTURE SCOPE

Most of the classification problems require optimization of model parameters along with multiple conflicting objectives such accuracy, specificity, mean squared error, precision, recall, misclassification rate etc. Multi-objective optimization is natural choice when dealing with classification problems. In literature, varieties of multi-objective evolutionary algorithms have been proposed for solving different classification problems. Also multi-objective optimization approach have been used for solving data mining tasks such as clustering, feature selection, association rule mining. Using this approach one can find set of optimal solutions for classification problems. As most of the multi-objective optimization approaches try to optimize two objectives so, we can consider more than two objectives for optimization. Also we can use multi-objective optimization for multi-class label classification problems.

References


AUTHORS

Seema Mane pursuing M.E. degree in Computer Engineering from Maharashtra Institute of Technology, Pune in Pune University and received B.E. degree in Computer Engineering from Amrutvahini College of Engineering, Sangamner in Pune University in 2012and her research interest includes Data Mining.

Mrs. Shilpa Sonawani is working as Assistant Professor at, MAEER’s Maharashtra Institute of Technology, Pune. She received her degrees, B.E. (Comp. Sci. & Engg.) from College of Engineering Badnera-Amravati in 1996, and M.Tech. (Comp. Sci. & Engg.) in 2008 from Guru Gobind Singh Indraprastha University, Delhi. She is a life member of ISTE, New Delhi. Her area of interest is Data mining, Multiobjective evolutionary algorithms.

Dr. Sachin R. Sakhare is currently working as Professor at Vishwakarma Institute of Information Technology, Pune. He did his B.E. (Comp. Sci. & Engg.) from College of Engineering Badnera-Amravati in 1996, M.E. (Comp. Sci. & Engg.) in 2006 and Ph.D. from S.G.B. Amravati University in 2014 in the faculty of Engineering & Technology in the area of Operating Systems. He is a life member of ISTE, New Delhi, and Member of IAENG Hong Kong.

Mrs. Pradnya Kulkarni is currently working as Assistant Professor at, MAEER’s Maharashtra Institute of Technology, Pune. She received her degrees, B.E. (Computer Engineering) from Dr. D. Y. Patil College of Engineering, Kolhapur in Shivaji University in 2001 and M.E. (CSE and IT) from Vishwakarma Institute of Technology, Pune in Savitribai Phule Pune University in 2008. Her area of interest is Data mining and Image Processing.