ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2016

Optimization Algorithms for Feature Selection in Classification: A Survey

S.Ruba Arockia Archana, Dr.M.S.Thanabal

M.E Student, Department of Computer Science and Engineering, PSNA College of Engineering and Technology, India

Associate Professor, Department of Computer Science and Engineering, PSNA college of Engineering and Technology, Dindigul, India

ABSTRACT: Classification problems have a large number of features in datasets, but not all them are useful for classification. Irrelevant and redundant features reduce the performance. These features may be considered as noisy. In order to solve this problem we perform a feature selection process. It is a preprocessing technique for solving classification problem. Feature Selection aims to choose relevant features to achieve better performance than using all features. Feature selection main objectives are maximizing the classification performance and minimize the features. Inorder to improve the classification accuracy a survey on various Optimization algorithms which uses feature selection for solving classification problem is done.

KEYWORDS*:* Feature selection, Optimization algorithms, Classification problem.

I. INTRODUCTION

Several applications of classification engross a large amount of information and large number of attributes. It is an significant job in machine learning and data mining which is used to classify each instance in dataset into groups based on the information described by the features. Irrelevant and redundant features are not useful for classification performance due to the large space known as curse of dimensionality. Since the data are collected for reasons other than mining the data there may be some irrelevant or redundant features. Many Classification tasks require learning of an appropriate classification task that assigns a given input to one of the restricted set of classes. The growth of the soaring throughput technologies lead to exponential growth in the harvested data with respect to both dimensionality and sample size. Dimensionality reduction is one of the well-liked techniques to remove noisy and redundant features .Feature selection is one of the dimensionality reduction technique used to remove irrelevant features and improve the classification accuracy.

II. RELATED WORK

In [4] author used Steganalysis it is a powerful tool to detect the presence of secretly hidden data in an object. The Steganalysis problem can be formulated as supervised classification problem and solved using machine learning approaches. The performance of machine learning approaches for Steganalysis is highly influenced by feature extraction and the nature of classification algorithm. In [7] author used Ant Colony Optimization (ACO) it models the problem as the search for a minimum cost path in a graph. Algorithms that tend to solve problems using ACO create a search space with nodes and design procedure to find a solution path . It generates the artificial ants and works in an iterative fashion to determine solution for the problems through their indirect communication via synthetic pheromone. In [8] author develop a filter feature selection algorithm based on PSO and the mutual information for each pair of features, which is used to evaluate the relevance and redundancy in the selected feature subset, and investigate whether this algorithm can select a small number of features to achieve better performance than using all features. It also develop a filter feature selection algorithm based on PSO and entropy of each group of features which is applied to evaluate the relevance and redundancy. In [9] author uses Particle Swarm Optimization (PSO) is population based evolutionary computation technique which simulates the behavior of organisms i.e. birds in a flock or fish in a school.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2016

Here each solution can be considered as individual particle in a given search space, which has its individual position and velocity. During movement each particle adjusts its position by changing its velocity based on its own experience as well as the experience of its experience until and optimum position is reached by itself and its neighbours.

III. CLASSIFICATION AND ITS APPLICATIONS

Classification is the problem identifying to which set of categories a new observation belongs on the basis of a training set of data contains observations whose category membership is known. Many real world problems can be modeled as classification problems such as assigning a specified email into spam or non spam classes. Assigning the categories of coming news as Sports and Entertainment and assigning a diagnosis to a given patient described by observed characteristics of patient like gender, blood pressure, presence or absence of certain symptoms.

IV. NEED FOR FEATURE SELECTION

Feature selection can address the curse of dimensionality by selecting only relevant features for classification. By eliminating and reducing irrelevant features and redundant features, feature selection could reduce the number of features, cut down the training time, simplify the learned classifiers and improve the classification performance. Incase of classification task the feature selection find negligible number of features so that the resulting probability distribution of the data classes with respect to the predicting attributes is as close as possible to that of the original dataset with all features.

A. A BINARY CUCKOO SEARCH AND ITS APPLICATION FOR FEATURE SELECTION[1]

The main idea of this paper is to use cuckoo search algorithm in the context of feature selection tasks .So binary version of the cuckoo search is proposed to evaluate it with different transfer functions that map constant solutions to binary solutions to binary ones. In addition, the Optimum Path Classifier accuracy is used as fitness function. We evaluate the strength of BCS to accomplish the feature selection task comparing it with the binary versions of Bat Algorithm, Firefly Algorithm, Particle Swarm Optimization .The provided simulations and analysis over four public datasets, employ traverse validation strategy to verify how the techniques work for feature selection purposes. The results confirmed that Cuckoo Search has good capabilities to find the best one on two out of four datasets.

B. MUTI-OBJECTIVEGENETIC ALGORITHM APPROACH TO FEATURESUBSET OPTIMIZATION[2]

The Setback of feature selection is multi-objective in nature and hence optimizing feature subsets with respect to any single evaluation criteria is not sufficient. Thus in order to combine several feature selection criteria the muti-objective optimization of quality subsets apply Multi –Objective Genetic Algorithm . The results confirm that the proposed system is able to determine diverse optimal feature subsets that are well spread in the overall feature space and the classification accuracy of the ensuing feature subsets is reasonably high. The diverse solutions satisfy reasonable levels of optimality with respect to predictive power, non-redundancy and cardinality of feature subsets. These solutions provide the users with several choices of feature subsets.

C. NEW APPROACH FOR FEATURE SELECTION BASED ON ROUGH SET AND BAT ALGORITHM[3]

A new selection technique based on rough-sets and bat algorithm is proposed. Bat Algorithm (BA) is attractive feature selection in which the bats will discover best feature combinations as they fly within feature subset space. Bat Algorithm does not need composite operators such as crossover and mutation it requires only primitive and simple mathematical operators, and is computationally inexpensive in terms both memory and runtime .A fitness incorporates both the classification accuracy and the number of selected features and hence balances the classification performance and reduction size. The used rough-set based fitness function ensures enhanced classification result keeping also minor feature size Rough set theory provides a arithmetical tool to find out data dependencies and reduce the number of features included in dataset by purely structural method. Rough sets have been a advantage tool for medical applications. The complete solution to detect nominal reducts is to produce all possible reducts and choose one with negligible cardinality which can be done by constructing kind of discernibility function from the dataset and simplifying it.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2016

D. A NEW FEATURE SELECTION ALGORITHM BASED ON BINARY ANT COLONY OPTIMIZATION[7]

A new feature selection algorithm based on Ant Colony Optimization (ACO) called Advanced Binary ACO (ABACO) is proposed. Features are treat as graph nodes to construct a graph model. In this graph each feature has two nodes, one for selecting the feature and other for deselecting. Ant Colony algorithm is used to select nodes while ants must visit all features. At the end of a tour, each ant has of a binary vector with the same length as the number of features where 1 imply selecting and 0 imply deselecting the corresponding feature. The experimental comparison verifies that algorithm has a good classification accuracy using a small feature set than another existing selection ACO –based feature selection method. ABACO permits ants to explore all features, but in most of ACO based algorithms searching among features is stopped until stopping criterion is met. The proposed algorithm has a strong search capability in the setback space and can efficiently find a minimal feature subset.

E. CUCKOO OPTIMIZATION ALGORITHM FOR FEATURE SELECTION IN HIGH-DIMENSIONAL DATASETS[6]

By the idea of Binary Cuckoo Optimization Algorithm (BCOA) and Information Theory a new filter feature selection method for classification problem is proposed. The proposed algorithm is based on BCOA and The mutual Information of each pair of features which determines the significance and redundancy in the fitness function of the proposed algorithm are used to further improve the performance in terms of number of features and classification accuracy. An Artificial Neural Network is employed to evaluate the classification accuracy of the selected subset on the test sets. In filter approach feature selection is done as and preprocessing procedure and the search process is independent of a learning algorithm. The goal was developing a new feature selection algorithm based on BCOA and two information measures namely entropy and mutual information. The result suggests the proposed approach reduces the number of features and improves the classification approach.

F. CHAOTIC MAPS IN BINARY PARTICLE SWARM OPTIMIZATION FOR FEATURE SELECTION[9]

Binary Particle Swarm Optimization (BPSO) is used for feature selection. Two Chaotic maps are embedded in binary particle swarm optimization a logistic map and tent map. The purpose of chaotic map is to determine inertia weight of BPSO. A Chaotic Binary Particle Swarm Optimization (CBPSO) is used a method to implement feature selection and k-nearest neighbor method with leave one out traverse validation serves as a classifier to evaluate the classification accuracies .Here a wrapper model for feature selection is adopted. The two chaotic maps show different dynamic behavior. The behavior affects the search capacity of CBPSO. The two different chaotic sequences for the inertia value are applied to the feature selection process. The fallout show that CBPSO with tent map obtained higher accuracy than CBPSO with a logistic map.

G. FILTER BASED BACKWARD ELIMINATION IN WRAPPER BASED PSO FOR FEATURE SELECTION IN CLASSIFICATION[5]

A new feature selection based on Particle Swarm Optimization (PSO) and a local search that mimics the typical backward elimination feature selection method is proposed. The proposed algorithm uses a wrapper based fitness function, i.e the classification inappropriate rate .The local search is performed only on the global best and uses a filter based measure which aims to take advantages of both filter and wrapper approaches. The goal of the approach to reduce the dimensionality of data while maintaining or increasing the classification accuracy using all features. The backward elimination is designed to be fast and effective by using a mutual information based filter measure and further utilizing the information given by the position value. It performs the search on small sub-groups of the features selected by the global best, rather than on the whole set of selected features. Since it takes the advantage of both filter and wrapper approaches which increase the classification accuracy.

H. FEATURE SELECTION METHOD USING GENETIC ALGORITHM FOR THE CLASSIFICATION OF SMALL AND HIGH DIMENSIONAL DATA[10]

An competent feature selection method that finding and selecting informative measures from small or high dimension data which maximum the classification accuracy use genetic algorithm to search out and identify the potential informative features combination for classification and then use the classification accuracy from the support vector machine classifier to decide the fitness in genetic algorithm. Chromosome representation is used in genetic algorithm.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 2, February 2016

A bit value of 1 in the chromosome means the feature is included in specified subset and bit value of 0 means the corresponding feature is not included in the specified subset. The main components of genetic algorithm are feature subset selection and SVM as classifier. The model of the chromosome representation in the proposed approach reduces the combination number of feature subsets with fitting chromosome length. The model further decreases the complexity searching on feature database. The results show the selected features are good to get high classification accuracy for training data of small or high dimension data.

V. CONCLUSION

This survey paper underlines the need for classification. The majority of the real world classification requires supervised learning. It is difficult to learn good classifiers before removing the unwanted features due to huge size of data. Reducing the irrelevant features can drastically reduce the running time of learning algorithms and yield good classifier. This arises the need of feature selection to remove irrelevant features.

REFERENCES

1. L.A.M.Pereira, D.Rodrigues, T.N.S.Almeida, C.O.Ramos ,A.N.Souza ,X.S.Yang and J.P.Papa, "A Binary Cuckoo Search and Its Application for feature Selection" ,Springer International Publishing Switerzland 2014.
2. Saroj,Jyoti, "Multi-Objective Genetic Algorithm Approach to Feature SubsetOptimization", IEEE International Advance Computing Conference(IACC) 2014.
3. E.Emary,W.Yaman, A.E.Hassanien "New approach for feature selection based on rough set and bat algorithm", Computer Engineering & Systems(ICCES)2014 9th International conference ,IEEE Conference Publications 2014.
4. Vasily Sachnev ,Hyoung Kim, "Binary Coded Genetic Algorithm With Ensemble Classifier for Feature in JPEG Steganalysis" , IEEE Nineth International Conference On Intelligent Sensors, Sensor Networks And Information Processing April 2014.
5. Hoai Bach Nguyen, Bing Xue, Ivy Liu ,Mengjie Zhang. "Filter Based Backward Elimination in Wrapper based PSO For Feature Selection Classification",IEEE Congress On Evolutionary Computation(CEC) July 2014.
6. Mahmood Moghadasian, SeyedehParvaneh Hosseini, "Binary Cuckoo Optimization Algorithm for Feature Selection in High –Dimensional Datasets" ,International conference on Innovative Engineering Technologies (ICIET)2014.
7. Shima Kashef,Hossein Nezamabadi "A new Feature Selection Algorithm based on binary ant colony optimization" ,IEEE 5th conference on Information and Knowledge Technology(IKT)2013.
8. Cervante L,Bing Xue,Mengjie Zhang,Lin Shang "Binary Particle Swarm Optimization for Feature Selection :A Filter Based Approach", IEEE Evolutionary Computation(CEC) 2012.
9. Cheng- SanYang, Li-Yeh Chuang, Jung- Chike Li,Cheng-Hong "Choatic Maps in Binary Particle Swarm Optimization for Feature Selection", IEEE Conference on Soft Computing in Industrial Applications June 2008.
10. Mohd Saberi Mohamad,Safaai Deris,Safie Mat Yatim,Muhammad Razib Othman "Feature Selection Method Using Genetic Algorithm For The Classification Of Small and High Dimensional Data", First International Symposium on Information and Communication Technologies October 2004.

BIOGRAPHY

S.Ruba Arockia Archana is a PG Scholar Studying in PSNA College of Engineering And Technology,Dindigul. She received Bachelor of Engineering(BE) degree in 2014 Anna University Chennai,India.

M.S.Thanabal is a professor in the Computer Science and Engineering Department,PSNA College of Engineering And Technology,Dindigul. He received Ph.D degree in 2015 form Anna University of Chennai,India.His research interests are Computer Networks,Data mining,Web mining.