

# Feature selection by convex optimization on UCI dataset survey

Pallavi Dixit<sup>1</sup>, Dr. Santosh Kumar<sup>2</sup>

Research Scholar<sup>1</sup>, Associate Professor<sup>2</sup>

Maharishi University of information Technology, Lucknow

**Abstract** - Feature selection process is mainly used in data preprocessing. This process reduce the high dimension data to low dimension data. Feature selection working as a feature weighting methods, we have to specify the number of selected features. it is often unknown what the optimal number of selected features is. This review analysis based on optimizes and normal method of features selection on structured, generic and heterogeneous data stream.

## I. INTRODUCTION

Feature selection is an approach of selecting subset from original features. Feature subset is measured by the means of evaluation criteria. Numbers of features are also increased when the dimensionality is increased. Feature selection process is mainly used in data mining for classification, clustering, association and regression. This process removes the irrelevant data and noise from the data set. Feature selection improves the quality of data and enhanced the processing speed of the algorithm.

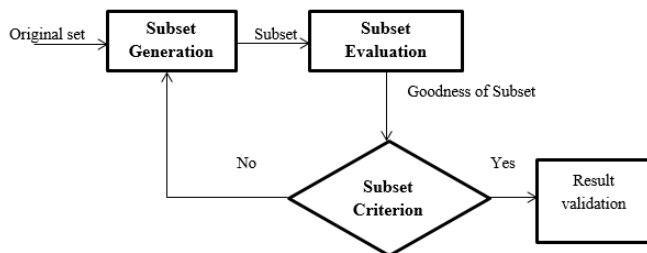


Fig.1: Features Selection Process

Steps that are followed in feature selection are:

1. Generation of Subset
2. Evaluation of Subset
3. Stopping Criterion
4. Validation of Results

**There are techniques used for the reduction of attributes:**

- 1) **Filter approach:** Filter approaches are not dependent of a learning technique and they are debated to be computationally inexpensive and more public than wrappers. It is used as a proxy measure rather than error

rate to score features. They are less computational as compare to wrapper approach.[1]

- Information Gain
- Chi-Square

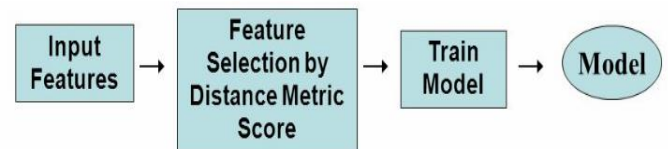


Fig.2: Filter Approach

- 2) **Wrapper approach:** It contains a learning technique as part of the evaluation function. Thus, wrappers can usually obtain best results from filter approaches. It is a predictive model to score subsets. Every individual subset is used to train a model which is then tested on a hold-out-test. They are very computational intensive but usually provide the best performing features set for particular type of model. [1]

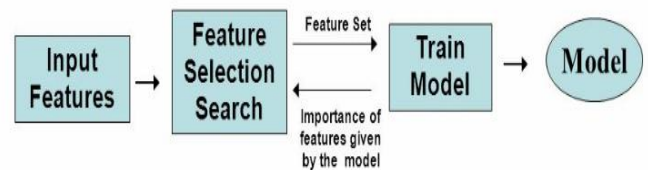


Fig.3: Wrapper Approach

- Forward selection
- Backward Elimination

- 3) **Embedded approach:** It is the combination of above two discussed method. A learning algorithm can take benefits of its own variable selection process and performs feature selection and classification simultaneously.

## II. RELATED STUDY

**A. Artificial bee colony algorithm:** Artificial bee colony algorithm is used for feature selection and it is integrated with genetic operators and sorting procedure. In this ABC algorithm is used in two ways that are ABC with continuous

and ABC with binary representation. The performance of the proposed system is compared with stepwise backward selection approach [1].

**B. Particle Swarm Optimization:** Particle swarm based feature selection approach is proposed by the author in this article. In this method, search is improved by using extra learning approach with subset cardinality. Learning on cardinality helps in position updates of swarms. Performance evaluation of the proposed method is done with the help of naïve Bayes and K-nearest neighbor algorithms [2]. The author proposed cost-based feature selection approach by using Multi-objective Particle Swarm Optimization. It generates the Pareto front of the non-dominated solutions that feature subset. Probability-based encoding is used to enhance the search capability. This method solves the cost based feature selection problem [3]. In this paper the author proposed to improve PSO and hybrid wrapper filter for feature selection. This approach reduces the computation complexity by providing effective features. This work is completed into two phase in the first phase two filters are applied to select features according to their ability and in the second phase reduce the irrelevant features according to their working. Particle swarm optimization is done in the first phase to get optimized results in the form of features. Performance evaluation is done on spatial domain and transforms domain [4]. The author proposed t-test feature selection approach for text categorization. In this paper the author mainly solve the issue of feature selection by constructing feature selection function based on term frequency. Diversity of the function is measured by using T-test. The results show the effectiveness then other approaches and performs better with KNN and SVM classifiers [5].

**C. Genetic Algorithm:** Genetic programming approach is proposed by the author in this paper for feature selection. This technique performs well on balanced and unbalanced data. It combines the feature sets which are selected by the distinct feature selection process. The proposed method is evaluated on the biological and textual datasets. It also increases the efficiency of learning process and enhances the effectiveness of the system [6]. In this paper the author proposed genetic algorithm feature selection approach to diagnosing the breast cancer. In this work, two different data sets are used for experimentation. The proposed work done in two stages. In the first stage, it eliminates the irrelevant features and genetic algorithm is used for extraction of useful information. This process reduces the computation complexity and enhances the speed of data mining. In second stage multiple classifiers are used to construct the effective model. Rotation forest method with G.A gives high accuracy in results [7]. Proposed the genetic algorithm for feature selection rainfall forecasting in sugarcane areas. In this paper, the author proposed genetic

algorithm feature selection approach to diagnosing the breast cancer. In this work, two different data sets are used for experimentation. The proposed work done in two stages. In the first stage, it eliminates the irrelevant features and genetic algorithm is used for extraction of useful information. This process reduces the computation complexity and enhances the speed of data mining [8]. The main objective of this paper is to identify the best algorithm in terms of more speed and accuracy in finding the solution, where speed is measured in terms of function evaluations. Achieving the goals basically depend on two key issues, less number of function evaluations (higher speed), or accuracy in approaching the answer. The author also described GA suffered from disadvantages such as high processing time and getting stuck in local minima and particle swarm optimization (PSO) algorithm easily to fall into local optimum in high-dimensional space and has a low convergence rate in the iterative process [9]

**D. Deep Learning Approach:** deep feature extraction and selection method which combines the weighted feature and stack feature selection together. The stack auto encoding method is helpful in the representation of information from the inputs. These features are combined with modified weighted features selection which is inspired by shallow structured machine learner. Set of features is used to reduce the bias of proposed learner model and also computational complexity [10]. Deep learning approach is used for feature selection and humanoid push recovery and classification. Four different types of pushes are used in this experiment. In this first classifier is based on the feed-forward back propagation neural network. And the second classifier is based on the deep neural network. The result validation is done on five-fold validation approach [11].

**E. Wolf Search Algorithm:** Author described that Heuristic optimization methods have an edge over their classical counterparts because they give globally optimum solution by using heuristics to efficiently search a large space. A special kind of heuristic optimization known as nature-inspired optimization or meta-heuristics is gaining popularity due to its advantages, which are applicable in computational intelligence, data mining, and their applications. Also described preying behavior of wolves and has displayed unique advantages in efficiency because each searching agent simultaneously performs autonomous solution searching and merging. The WSA's potential contributions to finding optimal solutions in applications include traveling salesman problems, quadratic assignment problems, and job scheduling problems [12]. Author proposed a new bio-inspired swarm intelligence algorithm named as Wolf search optimization for attribute reduced in classification. As we all know, Redundant and irrelevant attributes might minimize the classification accuracy because of the huge search space. Attribute reduction

techniques can be in general divided into two categories filter approaches and wrapper approaches. Wrapper approaches contain a learning technique as part of the evaluation function. Thus, wrappers can usually obtain best results from filter approaches. Forward selection in feature selection starts with an empty feature set (no features) and searches for a feature subset(s) with one feature by selecting the feature that achieves the highest classification performance. Backward selection starts with all the available features, then-candidate features are sequentially removed from the feature subset until the further removal of any feature does not increase the classification performance [13].

Sr. No	Algorithm	Observation
1	<b>Artificial bee colony algorithm</b>	<ul style="list-style-type: none"> <li>Forward selection and backward selection model are used.</li> </ul>
2	<b>Particle Swarm Optimization</b>	<ul style="list-style-type: none"> <li>KNN and SVM is used for classification of feature and PSO provides effective optimization.</li> </ul>
3	<b>Genetic Algorithm</b>	<ul style="list-style-type: none"> <li>Increases the efficiency of learning process and enhances the effectiveness of the system.</li> <li>reduces the computation complexity</li> </ul>
4	<b>Deep Learning Approach</b>	<ul style="list-style-type: none"> <li>Reduce the bias of proposed learner model.</li> <li>Classifier is based on the feed-forward back propagation neural network.</li> <li>Five-fold validation approach is used</li> </ul>
5	<b>Wolf Search Algorithm</b>	<ul style="list-style-type: none"> <li>Performance depends heavily on the manually chosen parameter values.</li> <li>Wolf search gives better result as compare other optimization algorithm.</li> <li>WSA is tested against classical algorithms such as GA and PSO</li> <li>It outperformed them in most of the testing cases.</li> </ul>

### III. CONCLUSION

Effectively preprocessing of data by reduction of dimensionality is called as feature selection. It is used for Machine learning application mean while used in statistics, patterns recognition and data mining. Feature selection method is used to reduce the dependency among the features. It reduces the dependency effect on reduction of parameters therefore it reduces the computation to finding the parameters. In this review, a brief description on the role of optimization approaches in features selection.

### IV. REFERENCES

- [1]. Hancer, Emrah, et al. "Pareto front feature selection based on artificial bee colony optimization." *Information Sciences* 422 (2018): 462-479.
- [2]. Hafiz, Faizal, et al. "A two-dimensional (2-D) learning framework for Particle Swarm based feature selection." *Pattern Recognition* 76 (2018): 416-433.

- [3]. Zhang, Yong, Dun-wei Gong, and Jian Cheng. "Multi-objective particle swarm optimization approach for cost-based feature selection in classification." *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)* 14.1 (2017): 64-75.
- [4]. Ghamisi, Pedram, Micael S. Couceiro, and Jon Atli Benediktsson. "A novel feature selection approach based on FODPSO and SVM." *IEEE Transactions on Geoscience and Remote Sensing* 53.5 (2015): 2935-2947.
- [5]. Viegas, Felipe, et al. "A Genetic Programming approach for feature selection in highly dimensional skewed data." *Neurocomputing* 273 (2018): 554-569.
- [6]. Aličković, Emina, and Abdulhamit Subasi. "Breast cancer diagnosis using GA feature selection and Rotation Forest." *Neural Computing and Applications* 28.4 (2017): 753-763.
- [7]. Haidar, Ali, and Brijesh Verma. "A genetic algorithm based feature selection approach for rainfall forecasting in sugarcane areas." *Computational Intelligence (SSCI), 2016 IEEE Symposium Series on. IEEE*, 2016.
- [8]. Shahla Shoghian, Maryam Kouzehgar "A Comparison among Wolf Pack Search and Four other Optimization Algorithms" *Vol 6 International Journal of Computer, Electrical, Automation, Control and Information Engineering*.
- [9]. Shahla Shoghian, Maryam Kouzehgar "A Comparison among Wolf Pack Search and Four other Optimization Algorithms" *Vol 6 International Journal of Computer, Electrical, Automation, Control and Information Engineering*.
- [10]. Aminanto, Muhamad Erza, et al. "Deep Abstraction and Weighted Feature Selection for Wi-Fi Impersonation Detection." *IEEE Transactions on Information Forensics and Security* 13.3 (2018): 621-636.
- [11]. Semwal, Vijay Bhaskar, Kaushik Mondal, and Gora Chand Nandi. "Robust and accurate feature selection for humanoid push recovery and classification: deep learning approach." *Neural Computing and Applications* 28.3 (2017): 565-574.
- [12]. Simon Fong, Suash Deb, Xin-She Yang "A heuristic optimization method inspired by wolf preying behavior" *Vol 26 Neural Computing & Application*, 2015.
- [13]. Wang, Deqing, et al. "t-Test feature selection approach based on term frequency for text categorization." *Pattern Recognition Letters* 45 (2014): 1-10.