Indian Journal of Science and Technology, *Vol 9(26), DOI: 10.17485/ijst/2016/v9i31/79656, August 2016*
ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

Optimized Feature Selection Algorithm for High Dimensional Data

D. Sheela Jeyarani[1*] and A. Pethalakshmi[2]

[1]Department Of Computer Science, Mother Teresa Women's University,Kodaikanal - 624101, Tamil Nadu, India; sheelajeyarani@gmail.com
[2]Department of Computer Science, M.V.M Government Arts College (W),Dindigul - 624001, Tamil Nadu, India; pethalakshmi@yahoo.com

Abstract

Objectives: This research paper, based on fuzzy entropy, adapts a new method along with firefly concept, seeks to select quality features. At the same time it removes redundant and irrelevant attributes in high dimensional data. **Methods/ Statistical Analysis**: Feature selection can be understood as a data prepossessing method in order to reduce dimensionality, eliminate irrelevant data and sharpening of accuracy. In the pattern space, fuzzy entropy is used to estimate the knowledge of pattern distribution. The study of the lightning quality of the fireflies has led to the introduction of the Firefly Algorithm for computing models. This work proposes an algorithm for selecting features by integrating fuzzy entropy and firefly algorithm. Our proposed algorithm's performances are analyzed using four different high dimensional data sets WILT, ORL, LC and LTG. **Findings:** The algorithm which is introduced here is further experimented with four variant data sets and the results shows that this algorithm out performs the traditional feature selection method. Also our proposed algorithm achieves maximum relevance and minimum level of redundancy. The performance metrics such as sensitivity, specificity and accuracy gives significant improvement when compared with existing FCBF algorithm. **Applications/Improvements:** Our optimized proposed algorithm efficiently improves the performance by eliminating redundant, noisy and insignificant features and can be applied on all high dimensional data sets.

Keywords: FCBF, Feature Selection Algorithm, Firefly Algorithm, Fuzzy Entropy, High Dimensional Data

1. Introduction

Our times have witnessed a phenomenal rise in the day to day collection and storage of huge data in various domains. There is an accumulation of vast data sets and it results in applications like machine learning, recognition of pattern, signal processing, text mining and web mining which has become very hazardous to develop or extort. In order to understand a given fact of concern, all the variables measured are not necessary in a high dimensional data. So attribute reduction becomes a major issue in handling high dimensional datasets. Data mining is vital in the course of discovery of knowledge in databases and is inevitable to derive important business information. It involves many stages like cleaning, reduction, integration, conversion, mining, pattern identification and presentation of knowledge. There are two types in data mining tasks namely, descriptive and predictive. What summarizes the relevant properties of data is termed as descriptive mining and what reveals the necessary information by surfing the present data is called as predictive mining. The various models of data mining are classification and prediction, clustering and association rules. This study concentrates on the applications of clustering and classification.

There are many stages to eliminate the redundant and irrelevant data from high dimensional data among which dimensionality reduction is an important one. One of the popular methods adopted for this purpose is feature selection. It is a method to select good features to achieve the target concept. There is an added advantage in mining on a minimized set of features. It decreases the features shown in the patterns identified in order to simplify to understand[1], Sequential search[2], non linear optimizations[3]

**Author for correspondence*

and Genetic Algorithms which are search approaches on feature selection algorithm. The feature selection process proceeds in the forward and the backward directions known as forward selection and backward selection method respectively. Forward selection begins without any variables and the variables are added at each step. It minimizes the error rate to the maximum and even to the extent of no further increase of variables does not reduce the error. On the contrary, backward selection begins with the entire set of variables. Consequentially, it eliminates the variables one after the other at every stage. It decreases the error rate to the maximum, even to the level of any further exclusion increases the error.

Fuzzy entropy enables partition of the input space into decision regions in order to choose appropriate features with good separability for the classification task. This idea has been defined in diverse ways[4–10] as well as so widespread in applied areas like economics, mathematics, communications and statistical thermodynamics[11–13]. Meta heuristic algorithm assumes a vital role in soft computing and computational intelligence in the sphere of present day global optimization algorithms. A subset of meta heuristics is called as 'Swarm Intelligence' (SI) based algorithm which are developed by imitating the behaviors of natural agents such as fish, birds, humans and others. Among the nature inspired optimization algorithms, Firefly Algorithm is one which can produce effective solutions for a large number of problems. This study presents a novel approach for selecting features by integrating fuzzy entropy and firefly algorithm.

2. Related Work

Machine Learning (ML) is an emerging branch of study with numerous applications in various fields like Natural Language Processing[14], Bio Informatics[15] and Image Processing[16]. It facilitates the analysis of large quantities by providing the necessary tools. One of the important basics in study of Machine Learning is feature selection. The feature selection identifies the important features, to enable the Machine Learning algorithm to concentrate on data which is helpful for researching and extracting new results. Algorithms on feature selection again and again identifies candidate subset and measures its optimality using the evaluation function. The feature selection approach, while it selects the candidate subset, intends to provide dimensionality reduction, insignificant data elimination, learning efficiency improvement and result accuracy[17,18]. The success of the feature selection depends on the dimensionality reduction of data and the improved performance of ML. Since the 1970s, feature selection has occupied a central place in the scientific researches, significantly in the area of identifying the statistical models[19,20] and Machine Learning [21].

Among the two approaches used to select features, namely, the filter approach and wrapper approach, the former is better in analyzing high dimensional data. Within the filter approach, there are various feature selection algorithms. It is broadly divided into two types, namely, Feature weighting algorithm and subset search algorithm. Relief[22] is very familiar among the feature weighing algorithms which weighs feature's capacity to distinguish events of various results depending on criteria related to distance. The disadvantage with Relief is that it is unable to deal with the removal of redundant features[23]. Relief-F[24], which is an extension of Relief, although is a competent algorithm in various ways, it is inefficient regarding the redundant features. As redundant features have adverse effects like increased computing time and lack of precision in learning algorithm, it has to be removed[25,26].

CFS[27], FCBF[28,29] and CMIM[30] are algorithms which deal with the redundant features. CFS emerges from the assumption that features in a good feature subset correlates extremely with the target class but does not correlate with each other. FCBF is an efficient fast filter method to identify relevant features and redundancy. CMIM repeatedly selects features to increase the mutual information. FAST[31] algorithm is distinct from the above mentioned algorithms and it selects features by using methods based on clustering.

The Genetic Algorithm[32] is one of the recent developments among the various types of feature selection algorithms. The GA is an evolutionary algorithm, biologically based and very efficient in scientific and engineering optimization[33]. Comparisons in the performance of GA with other classical algorithms has shown the usefulness of GA in the selection of features and had been published in several articles[34,35]. In GA, the chromosomes of the present generation is compared with fit chromosomes and selected credibly as seeds for the future generation by random mutation and cross over.

A new mathematical tool, rough set theory, deals with uncertainties and ambiguities of decision making. It had been actively used in all fields[36]. It selects the reduced set of set of all the qualities of the decision process. Medical data sets facilitate the introduction of rough set and neural net work based reduction[37]. Protocol based classification was proposed[38] using Genetic Algorithm with logistic regression and applied on KDD99 dataset. Methods to analyze data were formulated[39]. They deal with methods based on rules, Machine Learning and artificial intelligence.

A study deals with discernibility matrix and information gain to minimize features is presented[40]. Another finding introduces a new method about feature subset selection based on Fuzzy Entropy measures in making decisions[41]. Further an efficient feature selection technique using supervised Fuzzy information theory was proposed in[42]. In paper[43], offered a method which considers boundary samples while selecting features. Thus Fuzzy entropy method does not measure the purity of data sets. This method finds suitable attributes to obtain improved classification accuracy. A technique for innovative feature subset selection based on Fuzzy entropy measure was proposed in[44]. In[45] hybrid feature selection method is presented for network intrusion. In this paper, it deals with a new algorithm based on hybrid method which combines information gain and Genetic Algorithm in order to select features. Clustering is a technique used to choose features for classification. Paper[46] which gives an example of clustering technique that combines of two clustering algorithms to achieve a quality cluster of even small data points.

In handling modern global optimization problems[47–49], meta heuristic algorithms motivated by nature are very dominant. A subset of meta heuristics referred to as Swarm Intelligence (SI) based algorithms have been developed by imitating the natural behavior of biological agents such as birds, fishes, humans and others. The swarming behavior of birds and fishes led to determine particle swarm optimization[50], while the flashing pattern of tropical fireflies developed the firefly algorithm[51,52]. Likewise, the broad parasitism of some cuckoo species inspired the cuckoo search algorithm[53]. The Firefly Algorithm has proved to be very efficient among these new algorithms in dealing with multimodal global optimization problems.

Rough Set based Attribute Reduction (RSAR)[54] approach offers theoretical base to feature selection problems which need exhaustive search and which is not practical for large number of datasets. Various heuristic and random search strategies such as Quick Reduct and Entropy Based Reduct[55] are more efficient to overcome difficulties. Comparatively, algorithms inspired by nature are far more efficient to deal with complicated problems like optimization. Firefly OPF algorithm is implemented for finding optimal solution and sizing of STATCOM[56]. Several attempts have been made to bring together the RSAR approach with nature inspired algorithm to maximize the efficiency such as Gen RSAR[57], Ant RSAR[58], PSO-RSAR[59] and Bee RSAR[60] which are based on Genetic Algorithms, Ant colony optimization, particles swarm optimization and Bee colony optimization respectively.

However, the performance of Gen RSAR, Ant RSAR and PSO–RSAR are irregular and differ with the parameter values. Although Bee RSAR algorithm does not need any random parameter supposition, still it is slow to select a reduct[61]. Paper[62] presents a modern approach for feature selection based on nature inspired firefly algorithm. Papers[63,64] formulates a new meta heuristic algorithm by combining Levy' flights with a search strategy via firefly algorithm.

This study aims to propose *an* efficient algorithm for feature selection by integrating the fuzzy entropy and firefly algorithm. In Section 3, fuzzy entropy method, firefly algorithm, our proposed architecture and algorithms are discussed. Section 4 deals with the performance analysis of our proposed algorithm with existing works.

3. Methodology

3.1 Fuzzy Entropy based Feature Selection Method

Entropy is a quantity of uncertainty as a result of arbitrary research. The concept of fuzzy entropy is expansion to Shannon entropy in which the assessments of entropies are facilitated by the fuzzy sets. Fuzzy entropy is estimated by calculating the membership using Fuzzy C-Means clustering algorithm (FCM). Trivedi and Bezdeck presents Fuzzy C-Means algorithm given by,

$$L_m(U, v) = \sum_{k=1}^{n} \sum_{i=1}^{c} (U_k)^m \mid x_k - v_i \mid_A^2 P(x_j) \quad (1)$$

Where
$$u_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{\mid x_k - v_i \mid_A}{\mid x_k - v_j \mid_A} \right)^{2/m-1}} \forall i, k \quad (2)$$

$$vi = \frac{\sum_{k=1}^{n} (u_{ik})^m x_k}{\sum_{k=1}^{n} (u_{ik})^m} \quad (3)$$

Instead of each pattern, FCM algorithm computes the membership of every attribute in all clusters. The clusters are represented by centroid. Rather than calculating probability of a particular feature which will be in every precise class and then normalizing them, the feature's actual membership is found using FCM in the classes which also sum up to one. So the new term called match degree D_c is given by,

$$D_c = \frac{\sum_{xd \in c} U_c(x_d)}{\sum_{xd \in C} U_c(x_d)} \tag{4}$$

The fuzzy entropy of the elements of class c is given by,

$$FE_C = - D_C \log D_C \tag{5}$$

Fuzzy entropy along the universal set has to be calculated to find H(X),

$$\boldsymbol{E} = \sum_{c=1}^{C} \boldsymbol{E}_c \tag{6}$$

In order to eliminate those features which have no major contributions, a threshold r has to be fixed. Thus features are selected using fuzzy entropy based on threshold value.

3.2 Firefly Algorithm

Firefly Algorithm (FA) was introduced by Xin She Yang, based on the flashing manners of fireflies. The three major rules of fireflies are:

- Fireflies are unisex so that one is fascinated to the other despite of their sex.
- The attractiveness is proportional to their brightness and both (attractiveness and brightness) depends on the distance between fireflies.
- The landscape of objective function determines the brightness of the firefly. The less brighter firefly moves towards the brighter firefly. In case the less brighter firefly is unable to find a more brighter firefly, it moves in random.

The firefly's attractiveness is proportional to light intensity (β) seen by adjacent fireflies which varies with distance r and is given by $\beta = \beta_0 e^{-\gamma r^2}$ where β_0 is the attractiveness at r = 0. The movement of a firefly i is attracted to another more attractive (brighter) firefly j is determined by:

$$x_i^{t+1} = x_i^t + \beta_0 e^{-\gamma r_{ij}^2}(x_j^t - x_i^t) + atQ_i^t \tag{7}$$

where the second term is due to attraction. The third term is randomization with randomization parameter α_t and Q_i^t is a Gaussian distribution vector of random numbers. If $\beta_0 = 0$, it becomes a random walk. If $\gamma = 0$, it reduces to particles swarm optimization.

The basic Firefly Algorithm assumes that there exists n fireflies x_i (i = 1 . . . . n) initially placed arbitrarily in space. The light intensity I of each firefly is determined by the objective function f (x). ie I = α f (x). If $I_i > I_j$, j=1….n, j≠ i, the less brighter firefly j will move towards more brighter firefly i. The attractiveness varies with the distance $\gamma_{ij} = d(x_i, x_j)$. Also, the light intensity decrease with the distance from its source and it is also absorbed in the air determined by the absorption coefficient.

3.3 Proposed Work

This study aims to provide an optimized feature selection algorithm using Fuzzy entropy with Firefly concept in high dimensional data. Features can be extracted using fuzzy entropy and Fuzzy C-Means algorithm. The redundancy can also be removed. Firefly Algorithm is a powerful nature inspired optimization algorithm that can produce competitive solutions for a wide variety of problems. This work combines both fuzzy entropy and Firefly Algorithm. Figure 1 depicts the proposed architecture of this paper.

The following is the pseudo code of our proposed Optimized Feature Selection using Fuzzy Entropy and Firefly (OFEFLY) algorithm.

Procedure OFEFLY

{

Input: N – set of all features (f_i where i= 1 to N)

Output: S_{best} – Reduced set of features

- Begin,
- for i = 1 to N do begin,
- calculate $FE_{i,c}$ for f_i;
- if ($FE_{i,c}$ > = δ),

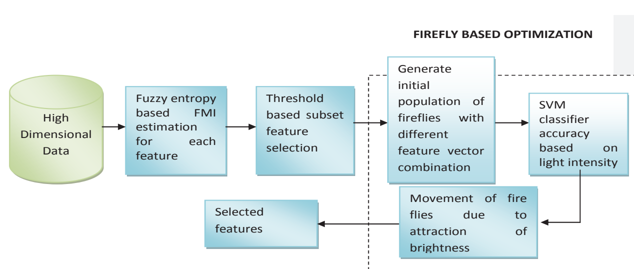

Figure 1. Proposed architecture for feature selection based on fuzzy entropy and firefly concept.

- append f_i to S'_{list}.
- end;
- Objective function f(x) : where f represents the SVM classifier accuracy; $X=(x_1,....x_d)^T$
- Generate initial population of fireflies x_i (i = 1 . . . . n) where $x_i \in$ subset of randomly selected features from S'_{list},
- Light intensity l_i at x_i is determined by $f(x_i)$,
- Define absorption Coefficient g,
- While (t< max Generation),
- for i = 1 : n all n fireflies,
- for j = 1 : i all n fireflies,
- if ($l_j > l_i$),
- Move firefly i towards j,
- end if,
- end j,
- end I,
- Rank the fireflies and find the current best $C_{best;}$
- end while,
- $S_{best} = C_{best,;}$
- end;

}

3.1 OFEFLY Algorithm

Given a data set with N features and a class c, the algorithm finds a set of predominant features S_{best} for the class concept. It consists of two parts. In the first part, it calculates fuzzy entropy value of each feature, selects relevant features into S'_{list} based on the predefined threshold δ. In the second part, it further calculates the light intensity using SVM classifier accuracy of each firefly according to the formula $l = l_0 e^{igr}$ where l_0 is the original light intensity and g is the fixed absorption coefficient. The attractiveness b of a firefly is proportional to the light intensity seen by adjacent fireflies and is given by $b = b_0 e^{igr^2}$ where b_0 is the attractiveness at r = 0. The movement of a firefly is attracted to another more attractive firefly j is determined by,

$$x_i^{t+1} = x_i^t + \beta_0 e^{-\gamma r_{ij}^2}(x_j^t - x_i^t) + \alpha t Q_i^t \tag{8}$$

If brightness of j is greater than i, move i to j. Thus convergence of firefly is determined by its movements.

4. Experimental Analysis

4.1 Experimental Datasets

In this section the performance of our OFEFLY algorithm is compared with the existing FCBF algorithm using four different data sets (WILT, ORL, LC and CTG). WILT data set contains some training and test data from a remote sensing study by Johnson et al. that involved detecting diseased trees in quick bird imagery. There are few training samples for the 'diseased trees' class (74) and many for 'other land cover class' (4265). ORL data set includes the database of faces contains set of face images taken between April 1992 and April 1994 at the lab. The database was used in the context of a face recognition project carried out in collaboration with the speech vision and robotics group of the Cambridge University Engineering Department. LC data set contains multivariate data set about lung cancer with number of instances 3200 and number of integer attributes 56. CTG (Cardio Tomography) is a monitoring of fetal heart rate and uterine contractions database consists in total of 552 intrapartum recordings with 21 attributes, which were acquired between April 2010 and August 2012 at the obstetrics ward of the University Hospital in Bruno, Czech Republic.

4.2 Result and Analysis

In this section, performances of FCBF and OFEFLY algorithms are analyzed. The algorithms are executed using MATLAB. All tests are done on Dual Core Processor with 2 GB RAM and 500 GB Hard Disk, running Windows 7. The analysis done with respect to Specificity, Sensitivity, Accuracy and Features selected are given below:

$$\text{sensitivity} = \frac{\text{number of true positives}}{\text{number of true positives} + \text{number of false negatives}}$$

$$= \frac{\text{number of true positives}}{\text{total number of sick individuals in population}}$$

$$= \text{probability of a positive test, given that the patient is ill}$$

$$\text{specificity} = \frac{\text{number of true negatives}}{\text{number of true negatives} + \text{number of false positives}}$$

$$= \frac{\text{number of true negatives}}{\text{total number of well individuals in population}}$$

$$= \text{probability of a negative test given that the patient is well}$$

Accuracy is the overall performance which is calculated using the function.

Accuracy = (TP+TN) / (TP+TN+FP+FN)

The findings of the entire four data sets with respect to Specificity, Sensitivity, Accuracy and Features selected are given in Table 1.

Figure 2 shows the comparative analysis of sensitivity ratio of our OFEFLY algorithm against FCBF algorithm.

Figure 3 shows the comparative analysis of specificity ratio of our OFEFLY algorithm against FCBF algorithm.

Figure 4 shows the comparative analysis of accuracy of our OFEFLY algorithm against FCBF algorithm and Figure 5 shows the comparitive analysis of feature selection process.

Table 1. Comparative analysis of OFEFLY with FCBF

		OFEFLY (FMI-FLYOP)	FCBF
WILT	Sensitivity	0.91042	0.8048
	Specificity	0.81242	0.69355
	Accuracy	0.85863	0.745045
	Features	40	60
ORL	Sensitivity	1	0.9126
	Specificity	1	0.9184
	Accuracy	1	0.915491
	Features	53.7037	75.92593
LC	Sensitivity	0.85714	0.7815
	Specificity	0.88889	0.80738
	Accuracy	0.875	0.794229
	Features	3.571429	41.6667
CTG	Sensitivity	0.84032	0.7652
	Specificity	0.91364	0.82247
	Accuracy	0.85137	0.792802
	Features	4.7619	33.3333

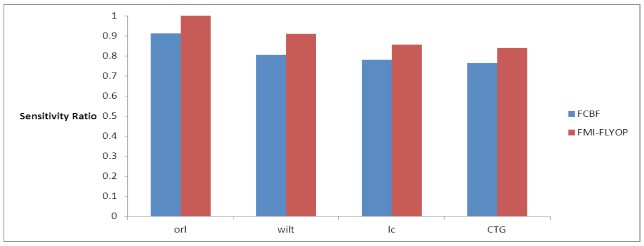

Figure 2. Sensitivity ratio analysis: OFEFLY vs. FCBF.

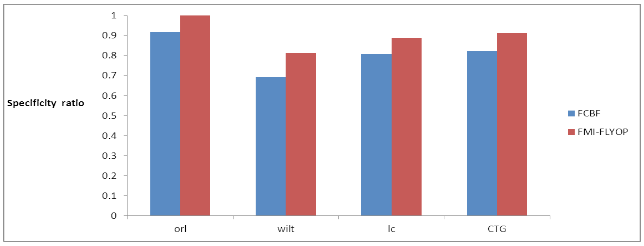

Figure 3. Specificity ratio analysis: OFEFLY (FMI-FLYOP) vs. FCBF.

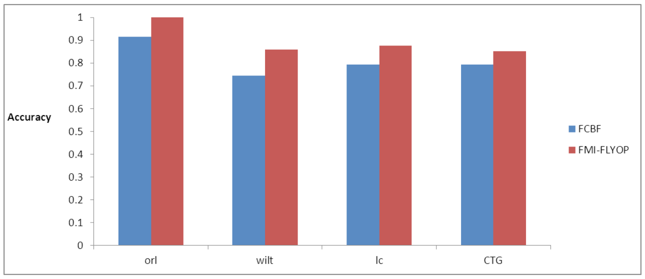

Figure 4. Accuracy analysis: OFEFLY (FMI-FLYOP) vs. FCBF.

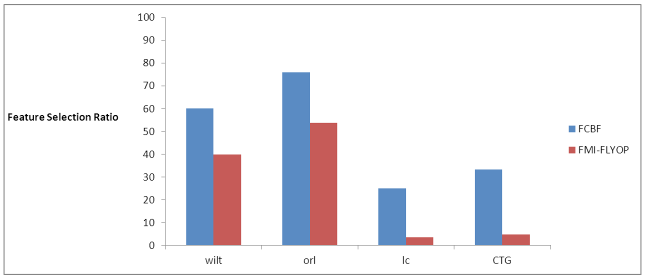

Figure 5. Feature selection analysis: OFEFLY (FMI-FLYOP) vs. FCBF.

5. Conclusion

The author introduced an optimized algorithm, utilizing fuzzy entropy and firefly concepts, for feature subset selection. This algorithm is developed to achieve maximum relevance and minimum degree of redundancy. Our OFEFLY algorithm performances are analyzed using four different data sets WILT, ORL, LC and CTG. We analyzed the performance metrics such as sensitivity, specificity, accuarcy and dimension metrics feature selection for all the data sets. Finally we compared the performances of OFEFLY and FCBF algorithm and proved that our proposed algorithm outperforms FCBF algorithm. Thus OFEFLY algorithm not only selects relevant features but also enhance the performance by removing redundant, meaningless, corrupted and insignificant features.

6. References

1. Han. J, Kamber M, Pei J. Data mining concepts and techniques. 2nd ed. USA: Morgan Kaufman Publishers; 2004.
2. Kittler J. Feature selection and extraction. Handbook of Pattern Recognition and Image Processing. Y. Fu, editor. New York: Academic Press; 1978.

3. Bradley PS, Mangasarian OL, Street WN. Feature selection via mathematical programming. INFORMS Journal on Computing. 1998; 10(2):209–17.
4. Oh IS, Lee JS, Moon BR. Hybrid Genetic Algorithms for feature selection. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2004; 26(11):1424–37.
5. Kosko B. Fuzzy entropy and conditioning. Information Sciences. 1986; 40(2):165–74.
6. Pal SK, Chakraborty B. Fuzzy set theoretic measure for automatic feature evaluation. IEEE Transactions on Systems, Man and Cybernetics. 1986; 16(5):754–60.
7. Shannon CE. A mathematical theory of communication. Mobile Computing and Communications Review. ACM SIGMOBILE. 2001; 5(1):3–55.
8. Hartley RV. Transmission of information. Bell Syst Tech J. 1928; 7:535–63.
9. Wiener N. Cybernetics. New York: Wiley; 1961.
10. Renyi A. On the measure of entropy and information. Proc Fourth Berkeley Symposium on Mathematical-Statistics Probability; Berkeley, CA. 1961. p. 541–61.
11. Belahut RE. Principles and practice of information theory. Reading, MA: Addison-Wesley; 1987.
12. Cover TM Thomas JA. Elements of information theory. New York: Wiley; 1992.
13. Ching JY, Andrew KC, Chan KCC. Class-dependent discretization for inductive learning from continuous and mixed-mode data. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1995; 17(7):641–51.
14. Blum AL, Langley P. Selection of relevant features and examples in Machine Learning. Artificial Intelligence. 1997; 97(1-2):245–71.
15. Baldi P, Brunak SA. Bioinformatics: The Machine Learning approach. Cambridge, Mass: MIT Press; 2001. p. 476.
16. Sajn L, Kukar M. Image processing and Machine Learning for fully automated probabilistic evaluation of medical images. Computer Methods and Programs Biomedicine. 2010; 104(3):75–86.
17. Dash M, Liu H. Feature selection for classifications: Intelligent data analysis. Intelligent Data Analysis. 1997; 1(3):131–56.
18. Kohavi R, John G. Wrappers for feature subset selection. Artificial Intelligence. 1997; 97(1-2):273–324.
19. Ben-Bassat M. Classification of pattern recognition and reduction of dimensionality. Handbook of Statistics; 1982. p. 773–91.
20. Siedlecki W, Sklansky J. On automatic feature selection. International Journal of Pattern Recognition and Artificial Intelligence. 1988; 2(2):197–220.
21. John GH, Kohavi R, Pfleger K. Irrelevant feature and the subset selection problem. Proceeding of the 11th International Conference on Machine Learning; San Francisco, CA: Morgan Kaufmann Publishers; 1994. p. 121–9.
22. Kira K, Rendell LA. The feature selection problem: Traditional methods and a new algorithm. Proc 10th Nat'l Conf Artificial Intelligence; 1992. p. 129–34.
23. Koller D, Sahami M. Toward optimal feature selection. Proc Int'l Conf Machine Learning; 1996. p. 284–92.
24. Kononenko I. Estimating attributes: Analysis and extensions of RELIEF. Proc European Conference on Machine Learning; Berlin Heidelberg: Springer; 1994. p. 171–82.
25. Kohavi R, John GH. Wrappers for feature subset selection. Artificial Intelligence. 1997; 97(1-2):273–324.
26. Hall MA. Correlation-based feature selection for discrete and numeric class Machine Learning. Proc 17th Int'l Conf Machine Learning; 2000. p. 359–66.
27. Hall MA. Correlation-based feature subset selection for Machine Learning. New Zealand: University of Waikato; 1999.
28. Yu L, Liu H. Feature selection for high-dimensional data: A fast correlation-based filter solution. Proceedings of the 20th Int'l Conference on Machine Learning. 2003; 20(2):1–8.
29. Fleuret F. Fast binary feature selection with conditional mutual information. Journal of Machine Learning Research. 2004; 5:1531–55.
30. Song Q, Ni J, Wang G. A fast clustering-based feature subset selection algorithm for high-dimensional data. IEEE Transactions on Knowledge and Data Engineering. 2013; 25(1):1–14.
31. Yu L, Liu H. Efficient feature selection via analysis of relevance and redundancy. Journal of Machine Learning Research. 2004; 5:1205–24.
32. Oh IS, Lee JS, Moon BR. Hybrid Genetic Algorithms for feature selection. IEEE Transactions on Pattern analysis and Machine Intelligence. 2004; 26(11):1424–37.
33. Das S, Ajith A, Konar A. Meta heuristic clustering. Berlin Heidelberg: Springer-Verlag; 2009.
34. Raymer ML, Punch WF, Goodman ED, Kuhn LA, Jain AK. Dimensionality reduction using Genetic Algorithms. IEEE Transactions on Evolutionary Computation. 2000; 4(2):164–71.
35. Yang JH, Honavar V. Feature subset selection using a Genetic Algorithm. Feature Extraction, Construction and Selection. New York: Springer Science; 1998. p. 117–36.
36. Yao YY, Zhao Y. Discernibility matrix simplification for constructing attribute reducts. Information Sciences. 2009; 179(5):867–82.
37. Thangavel K, Pethalakshmi A. Dimensionality reduction based on rough set theory. Applied Soft Computing. 2009; 9(1):1–12.
38. Yu KM, Wu MF, Wong WT. Protocol-based classification for intrusion detection. Applied Computer and Applied Computational Science (ACACOS '08). Hangzhou, China. 2008 Mar; 3(3):135–41.

39. Akbar S, Nageswara Rao K, Chandulal JA. Intrusion detection system methodologies based on data analysis. International Journal of Computer Applications. 2010 Aug; 5(2):10–20.
40. Azhagusundari B, Thanamani DAS. Feature selection based on fuzzy entropy. IJETTCS. 2013; 2(2):30–4.
41. Khushaba RN, Al-Jumaily A, Al-Ani A. Novel feature extraction method on fuzzy entropy and wavelet packet transform for myoelectric control. 2007 International Symposium on Communications and Information Technologies, ICSIT 07; Sydney, NSW. 2007. p. 352–7.
42. Sundary BA, Thanamani AS. An efficient feature selection technique using supervised fuzzy information theory. International Journal of Computer Applications. 2014; 85(19):40–5.
43. Shie JD, Chen SM. Feature subset selection based on fuzzy entropy measures for handling classification problems. Applied Intelligence. 2007; 28 (1):69–82.
44. Parvin H, Bidgoli B, Haffarin H. An innovative feature selection using fuzzy entropy. Advances in Neural Networks. 2011; 6677:576–85. ISSN 2011.
45. Sethuramalingam S, Naganathan ER. Hybrid feature selection for network intrusion detection. International Journal of Computer Science and Engineering. 2011; 3(5):1773–80.
46. Gayathri S, Mary Metilda M, Sanjai Babu S. A shared nearest neighbour density based clustering approach on a proclus method to cluster high dimensional data. Indian Journal of Science and Technology. 2015 Sep; 8(22). DOI: 17485/ijst/2015/8i22/79131.
47. Baeck T, Fogel DB, Michalewicz Z. Handbook of evolutionary computation. UK: IOP Publishing Ltd; 1997.
48. Bonabeau E, Dorigo M, Theraulaz G. Swarm Intelligence: From natural to artificial systems. New York: Oxford University Press; 1999.
49. Brown CT, Liebovitch LS, Glendon R. Levy flights in Dobe Ju/'hoansi foraging patterns. Human Ecology. 2007; 35(1):129–38.
50. Kennedy J, Eberhart R. Particle swarm optimization. Proc of the IEEE Int Conf on Neural Networks; Piscataway, NJ. 1995. p. 1942–8.
51. Yang XS. Nature-inspired metaheuristic algorithms. UK: Luniver Press; 2008.
52. Yang XS. Firefly Algorithms for multimodal optimization. Proc 5th Symposium on Stochastic Algorithms, Foundations and Applications. O. Watanabe and T. Zeugmann, editors. 2009. p. 169–78.
53. Yang XS, Deb S. Cuckoo search via Levy flights. Proceedings of World Congress on Nature and Biologically Inspired Computing (NaBIC); Coimbatore. 2009. p. 210–4.
54. Chouchoulas A, Shen Q. Rough set-aided keyword reduction for text categorization. Applied Artificial Intelligence. 2001; 15(9):843–73.
55. Dash M, Liu H. Feature selection for classification. Intelligent Data Analysis. 1997; 1(3):131–56.
56. Thilagar PP, Harikrishnan R. Application of intelligent Firefly Algorithm to solve OPF with STATCOM. Indian Journal of Science and Technology. 2015 Sep; 8(22). DOI: 10.17485/ijst/2015/v8i22/79100.
57. Liu H, Abraham A, Li Y. Nature inspired population based heuristics for rough set reduction. Rough Set Theory. SCI, Springer-Verlag. 2009; 174:261–78.
58. Jensen R, Shen Q. Finding rough set reducts with ant colony optimization. Proceedings UK Workshop on Computational Intelligence; 2003. pp. 15–228.
59. Yue B, Yao W, Abraham A, Liu H. A new rough set reduct algorithm based on particle swarm optimization. 2nd International Conference on the Interplay between Natural and Artificial Computation, IWINAC; 2007. p. 397–406.
60. Suguna N, Thanuskodi K. A novel rough set reduct algorithm for medical domain based on bee colony optimization. Journal of Computing. 2010; 2(6):49–54.
61. Trivedi M, Bezdeck JC. Low-level segmentation of aerial images with fuzzy clustering. IEEE Transactions on Systems, Man and Cybernetics. 1986; SMC-16(4):589–98.
62. Banati H, Bajaj M. Fire Fly based feature selection approach. IJCSI International Journal of Computer Science Issues. 2011; 8(4):473–80.
63. Yang XS. Firefly Algorithm, Levy flights and global optimization. Research and Development in Intelligent Systems. London: Springer; 2010. p. 209–18.
64. Yang XS He X. Firefly Algorithm: Recent advances and applications. International Journal of Swarm Intelligence. 2013; 1(1):36–50.