International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-2, July 2019

Accelerated Simulated Annealing and Mutation Operator Feature Selection method for Big Data

Renuka Devi D., Sasikala S.

Abstract: ***The optimal feature subset selection over very high dimensional data is a vital issue. Even though the optimal features are selected, the classification of those selected features becomes a key complicated task. In order to handle these problems, a novel, Accelerated Simulated Annealing and Mutation Operator (ASAMO) feature selection algorithm is suggested in this work. For solving the classification problem, the Fuzzy Minimal Consistent Class Subset Coverage (FMCCSC) problem is introduced. In FMCCSC, consistent subset is combined with the K-Nearest Neighbour (KNN) classifier known as FMCCSC-KNN classifier. The two data sets Dorothea and Madelon from UCI machine repository are experimented for optimal feature selection and classification. The experimental results substantiate the efficiency of proposed ASAMO with FMCCSC-KNN classifier compared to Particle Swarm Optimization (PSO) and Accelerated PSO feature selection algorithms.***

Index Terms: *Accelerated Simulated Annealing and Mutation Operator (ASAMO), Big Data, Feature selection, Fuzzy Minimal Consistent Class Subset Coverage (FMCCSC), K-Nearest Neighbor (KNN) classifier, Swarm Intelligence*

I. INTRODUCTION

The process of extracting meaningful information from voluminous data has been a demanding concern in Data mining and machine learning (ML) [7]. This context of the issue is commonly handled in "big data"[13]. It has involved with allied domains such as Bioinformatics, medicine, marketing, and finance [14]. The modern advancement of cloud computing technologies is useful in adapting the big data mining techniques over the massive amounts of data [8-9]. In classification method, Feature Selection (FS) is intended to select only the relevant and most influential features by removing the irrelevant and superfluous features, with enhanced accuracy and increased classification model building. Even though, comprehensive computing method may be employed for optimal feature subset selection, but is not the same in handling the high data streams which are gathered at a rapid rate. Amongst the various techniques, evolutionary methods have been successfully used for FS [19]. The excessive addition of the feature space leads to time complexity. Evolutionary Computation (EC) methods are fascinated towards nature stimulated techniques such as Genetic Algorithms (GAs), Genetic Programming (GP), Ant Colony Optimization (ACO), and Particle Swarm Optimization (PSO) algorithms. The evolutionary approaches are appropriate for the optimal FS by means of either wrapper or filter algorithms. The advantages of heuristic algorithms are search capability in selecting the optimal feature. Here, the various EC methods used in classification problems are reviewed [10]. The review enables us to expose the best EC algorithms for optimal feature selection and the future research directions in FS. Sadeg et al [15] developed Bee Swarm Optimization (BSO) algorithm for FS. The proposed algorithm is wrapper based FS with classifier algorithm. The experimental results confirm the improved performance of BSO in selecting the optimal features with improved classification accuracy. Fong et al [12] proposed an adaptable FS approach called Swarm Search Feature Selection (SS-FS). It is designed to handle higher dimensionality feature selection problems. SS-FS is confirmed to be a reasonable FS method with higher accuracy in classification and it is tested with biomedical data sets. Tennant et al [17] highlight the Micro-Cluster Nearest Neighbour (MC-NN) data stream classifier. The MC-NN is based on incremental approach. Here the data streams are added incrementally and processed without the requirement of residing in memory. Fong et al [16] proposed a novel, lightweight Accelerated Particle Swarm Optimization (APSO) feature selection algorithm for big data streams. APSO is based on the swarm intelligence and proved that the algorithm performed well in terms of accuracy, time complexity, and so on. However, the classification of the big data streams becomes a very difficult task. The key motivation of this work is how to handle big data streams efficiently. From this influence, Swarm Search with Accelerated Simulated Annealing and Mutation Operator (ASAMO) feature selection algorithm is applied to big data streams. The aim of the proposed ASAMO algorithm is to have a combination of classification as well as FS algorithm as one. The review of the feature selection methods for handling data stream is also discussed in the recent work [18-21].

Revised Manuscript Received on 30 July 2019.

* Correspondence Author
D.Renuka Devi*, Department of Computer Science, IDE,University of Madras, Chennai, India.
Dr. Sasikala. S, Department of Computer Science, IDE,University of Madras, Chennai, India.

II. PROPOSED METHODOLOGY

Feature selection is to eliminate irrelevant and redundant features, thus increasing the classification accuracy.

However, conventional methods fall short of scalability to handle with huge datasets in a delimited time. The novel ASAMO-FS method is proposed for mining big data. The outcome of the proposed experiment with the benchmark datasets has verified the efficacy of the algorithm in terms of performance metrics and results is compared to FS with PSO and APSO algorithm. Even though the features are reduced from feature subset; still the identification of classes becomes a major intricate task. In this work, we present a novel, Fuzzy Minimal Consistent Class Subset Coverage (FMCCSC) for computing a consistent subset of the K Nearest Neighbour (KNN) decision. The other conventional classifiers such as Support Vector Machine (SVM), random forest, Naïve Bayes, Hoeffding Tree (HT) are also experimented along with the FMCCSC-KNN classifier for classification.

A. *Architecture Design*

The most important features are selected by the ASAMO algorithm. The Fuzzy Minimal Consistent Class Subset Coverage (FMCCSC) is used to find the subset of the class's instances. Finally, present a novel FMCCSC-KNN for computing training set with the nearest neighbour decision rule. The previous steps are repeated until it reaches the maximum number of iterations, and countable instance subsets are obtained in the classification. The proposed model is shown in figure 1.

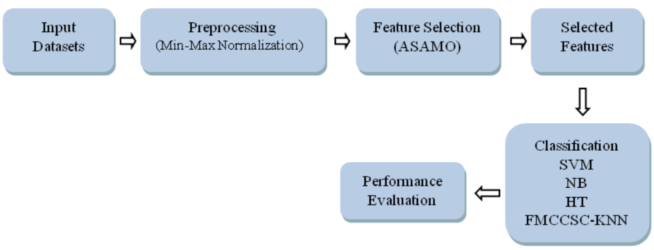

Fig.1. Proposed Model

B. *Preprocessing*

The datasets are preprocessed by min-max normalization method. It is the method of scaling the given dataset within the specified range of values between 0 and 1. Transform the data from measured units to a new interval from new_min $_A$ to new_max $_A$.

$$v' = \frac{v - min_A}{max_A - min_A}(new\_max_A - new\_min_A) + new\_min_A \quad (1)$$

C. *Accelerated Simulated Annealing and Mutation Operator (ASAMO)*

The proposed FS algorithm is to select an optimal feature subset from a big dataset. The ASAMO based FS model which maintains the classification accuracy and selects the optimal feature with highest fitness function value. The ASAMO Model is shown in Figure 2.

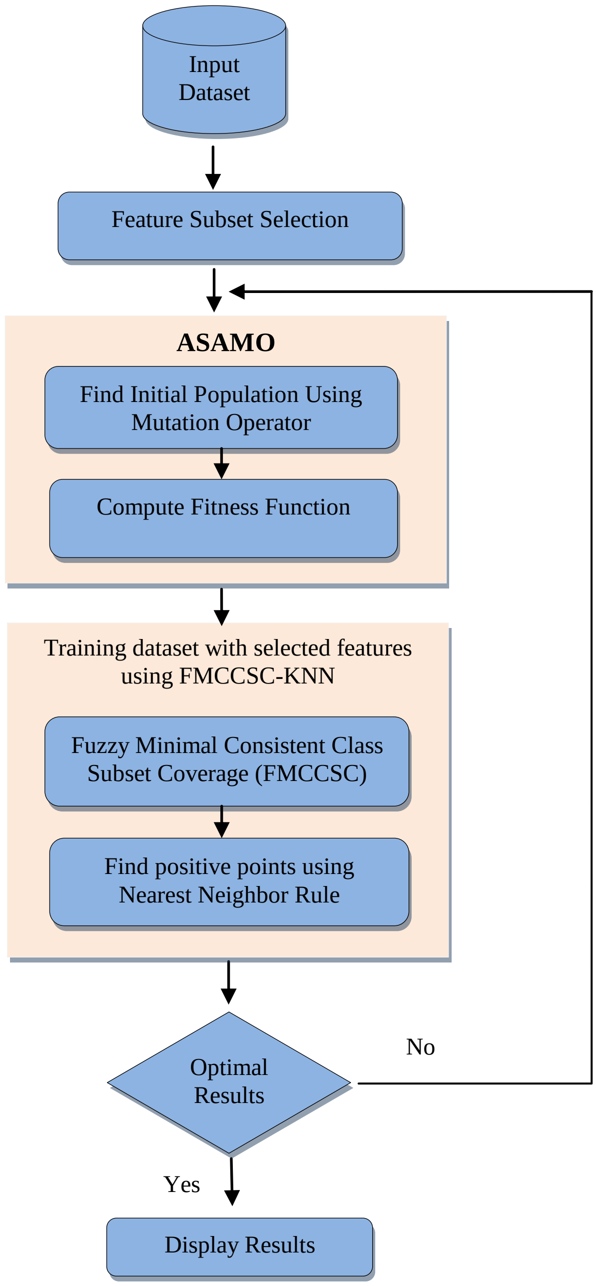

Fig.2.ASAMO Model

The Feature selection starts with a randomly searching the feature space and selecting the optimal feature which increases the classification accuracy in a stochastic manner.

D. *Simulated Annealing (SA)*

Simulated Annealing (SA) is the method of heating any metal or glass above its melting point, controlling the temperature for a definite time, followed by cooling it with a slower rate to solidify the metal into the perfect structure. This process produces the highest-quality structure of the metals. This simulation process is recognized as SA [1-2]. There is a correlation between the SA processes with feature optimization. The different states of metal correspond to potential solutions of the problem space. SA [2] is an alternative to the Metropolis algorithm, where the features are selected by changing the temperature level consequently maintaining the accuracy.

The two component parts of SA are as follows: one component method for the generation of optimal FS and the other for the approval of FS solutions. SA reduces the global minimum error value for big dataset classification. The optimal FS by this method produces the result with minimal error rate within the computational time. The parallel SA algorithms have been proposed for optimal FS with enhanced classification accuracy [5].

The system with energy E_α at temperature T, with the probability P_α is defined in (2).

$$P_\alpha = \frac{1}{Z} e^{-\frac{E_\alpha}{k_B} T} \tag{2}$$

Where ,

k_B = Boltzmann's constant

T = Absolute temperature is computed based on the fitness value of the classifier

$$Z = \sum_\beta e^{-\frac{E_\alpha}{k_B} T} \tag{3}$$

When T is higher, the classification accuracy is high. The Boltzmann function is distributed uniformly across different states, at any rate of the energy. When T reaches zero, the classification accuracy is lesser for the selected features. In (3), the constant k_B is ignored. At higher temperature, SA will search for optimal features with lesser results, ignoring the changes in the energy. When the temperature T is lowered, the feature is selected in the neighborhood of already selected features with minimum error rate and finds even better features with a minimum error value. When the temperature T reaches 0, it is the equilibrium state. When the state shifts for a time T, the probability of the change is determined by the Boltzmann distribution.

$$P = e^{-\frac{\Delta E}{T}} \tag{4}$$

The energy function ΔE, is inversely proportional to T.

Algorithm 1. Simulated Annealing based feature selection

1. *Initialize number of features as F=(f_1,...,f_m)*
2. *Randomize x(0)*
3. *Compute fitness function as classification accuracy and compute temperature*
4. *Repeat for (i=1…m)*

4.1. Repeat

a. State x is incremented by Δx

b. Compute ΔE(x)=E(x+ Δx)-E(x)

If ΔE(x)<0, maintain the state

or else

Accept the newly selected features with P

c. Decrement T by ΔT

d. Until T achieves higher accuracy

5. *End for*
6. *End*

The initial position of the SA is rearranged by using the Gaussian Mutation operator., to increase the classification accuracy. However, there is several numbers of mutation operators are presented in the literatures. But among them Gaussian Mutation operator is performed based on the probability distribution so it finds the accurate value of the specific variable. Let f_i [a_i, b_i] be a real variable. The probability distribution function using Gaussian Mutation operator is as follows,

$$p(f_i^{'}, f_i, \sigma_i) = \begin{cases} \dfrac{\frac{1}{\sigma_i}\phi(\frac{f_i^{'} - f_i}{\sigma_i})}{\phi(\frac{b_i - f_i}{\sigma_i}) - \phi(\frac{a_i - f_i}{\sigma_i})} if a_i \le f_i^{'} \le b \\ 0 otherwise \end{cases} \tag{5}$$

The mutation strength parameter σ_i ranges between the bounds a_i and b_i, $\sigma = \sigma_i/(b_i - a_i)$ as a fixed non dimensionalized parameter for all features(m). Where $\varphi()$ is the probability distribution of the standard normal distribution and $\Phi(\cdot)$ is the cumulative distribution function. Compute $f_i^{'}$,

$$f_i^{'} = f_i + \sqrt{2\sigma}(b_i - a_i) erf^{-1}(u_i^{'}) \tag{6}$$

$$u_i^{'} = \begin{cases} 2u_L(1 - 2u_i) if u_i \le 0.5 \\ 2u_R(2u_i - 1) if u_i \ge 0.5 \end{cases} \tag{7}$$

$$erf^{-1}(u_i^{'}) = sign(u_i^{'}) \left(\sqrt{(\frac{2}{\Pi\alpha} + \frac{\ln(1-u_i^2)}{2})^2 - \frac{\ln(1-u_i^2)}{2}} - \frac{2}{\Pi\alpha} + \frac{\ln(1-u_i^2)}{2} \right)^{1/2} \tag{8}$$

Where $\alpha = \frac{8(\Pi - 3)}{3\Pi(4-\Pi)} = 0.140012$ and sign($u_i^{'}$ <0) is -1 and $u_i^{'}$ <0 and is +1 if $u_i^{'} \ge$ 0 Also, u_L and u_R are calculated as follows

$$u_L = 0.5\left(erf\left(\frac{a_i - f_i}{\sqrt{2}(b_i - a_i)\sigma} \right) + 1 \right) \tag{9}$$

$$u_R = 0.5\left(erf\left(\frac{b_i - f_i}{\sqrt{2}(b_i - a_i)\sigma} \right) + 1 \right) \tag{10}$$

Hence, to mutate i^{th} feature variable, f_i is as follows:

Generate u_i, a random number between 0 and 1

Create offspring $f_i^{'}$ from the parent f_i

E. ***ASAMO***

The modified working principle of the proposed Accelerated Simulated Annealing and Mutation Operator (ASAMO) algorithm is discussed as follows:

Algorithm 2. ASAMO algorithm based feature selection

1. *Initialize number of features as $F=(f_{1},\ldots.f_{m})$*
2. *Randomize x(0)*
3. *Initialize new positions to create offspring $f_i^{'}$ from parent f_i*
4. *Compute fitness function from classification accuracy and compute temperature T*
5. *Repeat for (i=1…m)*

5.1. Repeat

e. State x is incremented by Δx

a. $\Delta E(x)=E(x+\Delta x)-E(x)$

If $\Delta E(x)<0$, maintain the state

Or else

Accept the newly selected features with P

b. Decrement T by ΔT

c. Until T achieves higher accuracy

6. *End for*
7. *End*

F. Fuzzy Minimal Consistent Class Subset Coverage (FMCCSC)

In this section, first, Minimal Consistent Class Subset Coverage (MCCSC) [6] is introduced; next, Fuzzy Minimal Consistent Class Subset Coverage (FMCCSC) problem is discussed. The subsets are considered to be consistent, when it satisfies the class constraint Nd . For a given known number of training dataset samples Xij, j=1 to m selected features, with the condition t=nearest neighbor point. This is represented by (Xij, Nd). Let C, consists of minimal number of subsets that satisfies the given constraint, the nearest neighbor point t from K Nearest Neighbor Algorithm (KNN).

G. Fuzzy Minimal Consistent Class Subset Coverage (FMCCSC) with KNN classifier

K-Nearest Neighbors (KNN) is commonly used for classification [3-4]. Let us consider k is the number of training samples. For exactly finding the nearest class value, KNN classifier is used which separates the class labels correctly. In a sample space $R^d \times \{1, 2\}$, there are n number of samples with output class Y_{ij}, so that $X_{ij}| Y_{ij}= r\tilde{}P_r$ and r=1,2. Given some norm ‖.‖ on R^d and a point $FX_{ij}\varepsilon R^d$, let (FX_{ij} , Y_{ij}) can be rearranged such that$\|FX_{1j}-FX\|<=.....<=\|FX_{ij}-FX\|$. During the training phase of the algorithm, it maintains the feature vectors belong to the specific class label. In classification phase, unlabeled vector is classified to the most frequent nearest class label of training samples. The distance between the points is calculated by Euclidean distance.

III. EXPERIMENT AND RESULTS

The proposed methodology is implemented with Dorothea and Madelon datasets (UCI machine learning repository). The Dorothea dataset is a drug discovery dataset which classifies either the drug is active or inactive. It consists of 1950 instances and 100000 features. The Madelon is an artificial dataset consists of 4400 instances and 500 attributes. The performance results of ASAMO are compared with PSO and APSO. The results of the Dorothea and Madelon datasets are tabulated in Table 1 and Table 2. The four classifiers are applied after feature selection and compared with other FS algorithms.

Table 1. Classification results of Dorothea

Feature selection /classifier		Precision	Recall	F-Measure	Accuracy	Processing time (seconds)
NB	FS-PSO	81.235	81.891	81.563	80.892	32.63
	FS-APSO	83.874	84.561	84.217	84.171	28.92
	FS-ASAMO	88.583	87.814	88.198	88.025	25.63
SVM	FS-PSO	76.814	77.812	77.313	78.253	52.81
	FS-APSO	77.581	79.561	78.571	79.581	45.89
	FS-ASAMO	81.258	82.171	81.714	82.589	41.26
HT	FS-PSO	84.124	85.512	84.818	85.814	21.81
	FS-APSO	86.891	87.785	87.338	86.715	20.71
	FS-ASAMO	89.975	90.171	90.073	90.523	18.58
FMCCSC-KNN	FS-PSO	87.215	88.512	87.863	89.154	18.52
	FS-APSO	89.281	90.752	90.016	91.512	17.41
	FS-ASAMO	91.584	92.638	92.111	92.891	15.81

The precision, recall, F-measure, and classification accuracy are the performance metrics used to evaluate the performance of this research work. Figure 3 shows the performance comparison results of different classifiers in terms of accuracy of Dorothea dataset. The accuracy metric signifies the precise classification of given datasets. Dorothea is the drug dataset, where a new drug is discovered by means of isolating the molecules into active and inactive compounds. Such a finding can lead to designing new compounds with the desired properties. The important features are selected by the proposed ASAMO and classification of drug into active and inactive is done by FMCCSC with KNN classifier. From the figure, the proposed FMCCSC-KNN classifier with three FS-PSO, FS-APSO, and FS-ASAMO feature selection algorithms produces classification accuracy results of 89.154%, 91.512%, and 92.89 %.

OPEN ACCESS
International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-2, July 2019

Table 2. Classification results of Madelon

Feature selection /classifier		Precision	Recall	F-Measure	Accuracy	Processing time (seconds)
NB	FS-PSO	64.148	75.878	70.013	72.847	44.56
	FS-APSO	66.666	86.956	75.471	74	42.89
	FS-ASAMO	70.156	71.255	70.705	75.154	39.02
SVM	FS-PSO	69.251	77.015	73.133	74.581	60.81
	FS-APSO	72.814	79.51	76.162	76.891	56.89
	FS-ASAMO	76.156	77.255	76.7055	78.715	50.15
HT	FS-PSO	79.175	79.171	79.173	80.521	28.92
	FS-APSO	80.769	80.769	80.769	80	26.52
	FS-ASAMO	84.615	81.481	83.018	82	21.56
FMCCSC-KNN	FS-PSO	89.258	88.814	89.036	88.015	23.62
	FS-APSO	91.251	91.058	91.154	91.812	21.58
	FS-ASAMO	93.581	94.152	93.866	94.25	18.91

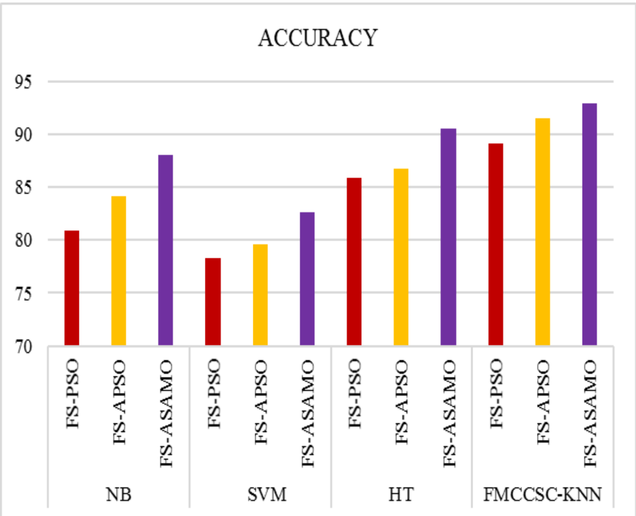

Fig.3. Accuracy Results Vs Classifiers

The precision metric explores the exactness of the proposed model through analyzing the predicted false positive values. Higher the precision value, lowers the false positive, thus lead to more precise model. This is shown in the Fig.4, for Madelon dataset.

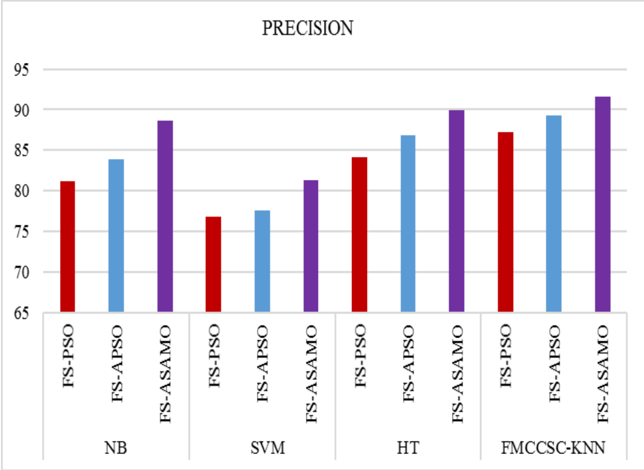

Fig.4. Precision Results Vs Classifiers

Recall measures are related to false negative values. In Dorothea database, if any, drug compound is falsely identified into inactive, thus affecting the drug formulation. Figure 5, shows the comparison.

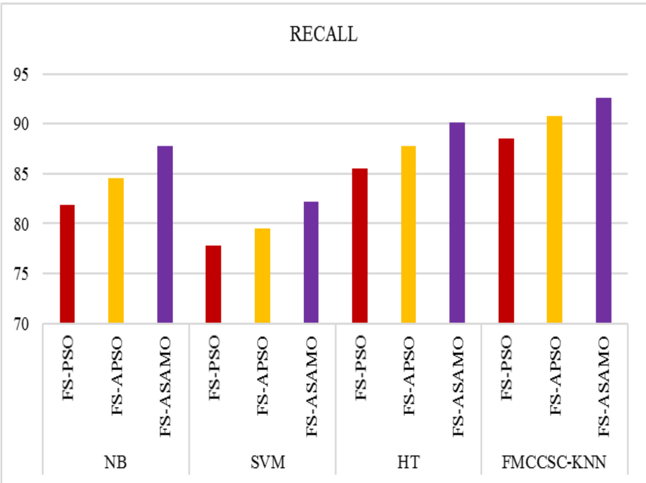

Fig.5. Recall Results Vs Classifiers

The F - score value identifies the accuracy of the model. It is a ratio of precision and recall. In Figure 6, demonstrates the results with the classifiers and other FS algorithms. It is a testimony of the ASAMO efficacy.

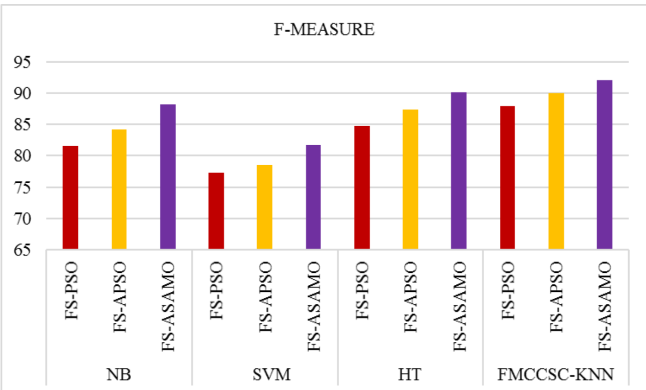

Fig.6. F-Measure Results Vs Classifiers

Figure 7-10, shows the performance comparison results of different classifiers in terms of metrics evaluated for Madelon dataset, which is an artificial dataset. Based on the features, we have to separate the examples into positive and negative class. The important features are selected by the proposed ASAMO and classification of into positive and negative class is done by FMCCSC with KNN classifier. From the figures, the proposed FMCCSC-KNN classifier with three FS-PSO, FS-APSO and FS-ASAMO feature selection algorithms produces classification accuracy results of 88.015%, 91.812%, and 94.25%.

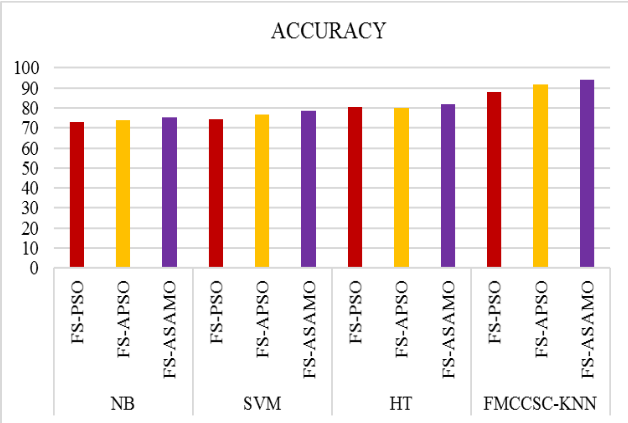

Fig.7. Accuracy results Vs classifiers

It demonstrates that proposed work performs better when compared to other classifiers. It demonstrates that proposed work performs better when compared to other classifiers. In the proposed work optimal features are selected by using the ASAMO and in the classification stage FMCCSC is introduced to find accurate class separation. These two parts increase the efficiency of the classifier.

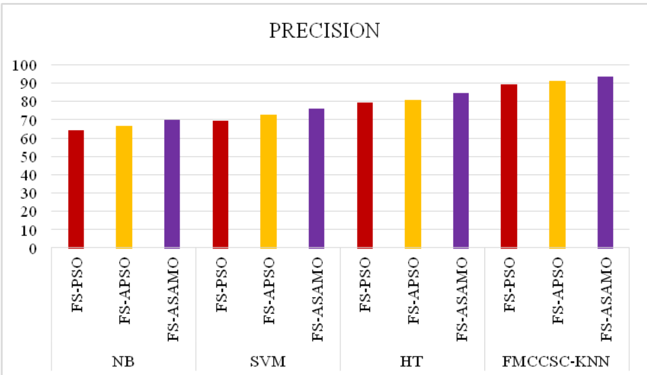

Fig.8. Precision Results Vs classifiers

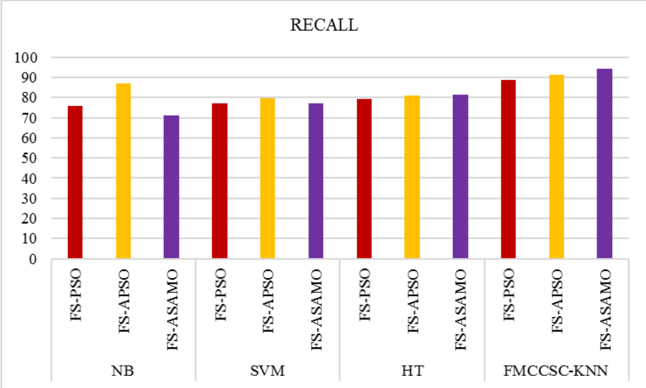

Fig.9. Recall Results Vs Classifiers

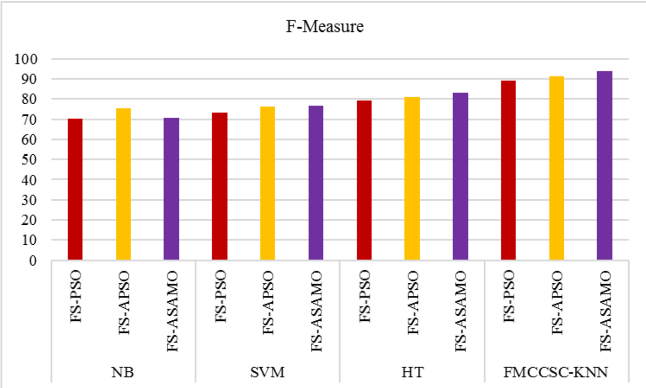

Fig.10. F-Measure Results Vs Classifiers

The ROC curve of the proposed algorithm is shown in the Figure 11.

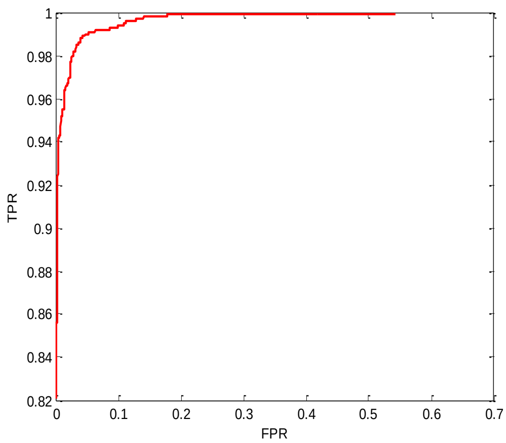

Fig.11. ROC of ASAMO Algorithm

IV. CONCLUSION

In this work, novel lightweight FS approach ASAMO is proposed for big data mining. The proposed method is experimented on voluminous data sets with an outsized quantity of features. The optimal features are selected by ASAMO and then classified into active and inactive compounds of drug discovery by the FMCCSC - KNN classifier in the case of Dorothea dataset. Similarly, in Madelon datasets, optimal features are selected and classified into positive or negative class. The proposed ASAMO feature selection method is evaluated by using three well-known classifiers (SVM, NBs, HT, and FMCCSC-KNN) and they are compared with PSO and APSO algorithm for FS. The outcome of the experiments is ascertained by the performance metrics like Accuracy, F-measure, Recall, and Precision. The efficacy of proposed algorithm is proven by the results and further compared with other state- of- the- art FS algorithms. The future challenges include parallel processing, as execution time is vital for any method. The execution time is optimal for the proposed work, however, selecting features from an exceptionally high dimensional data is really a challenge.

REFERENCES

1. Van Laarhoven PJ,Aarts EH. Simulated annealing. In simulated annealing: Theory and applications.Springer Netherlands.1987; 7-15.
2. Szu H, Hartley R. Fast simulated annealing. Physics letters A. 1987; 122(3-4): 157-162.
3. Song Z, Roussopoulos N. K-nearest neighbor search for moving query point. In International Symposium on Spatial and Temporal Databases. 2001;79-96.
4. Franco-Lopez H, Ek A R, Bauer ME.Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbors method. Remote sensing of environment. 2001; 77(3): 251-274.
5. Czech ZJ. Three parallel algorithms for simulated annealing. In: Proceedings of the 4th international conference on parallel processing and applied mathematics, Naczow, Poland. London: Springer; 2001;210–217.
6. Gao, BJ., Ester M, Cai JY,Schulte O, Xiong H. The minimum consistent subset cover problem and its applications in data mining. In Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining. 2007;310-319.
7. Alpaydin E.Introduction to Machine Learning. 2nd edition .MIT Press, Cambridge: Mass, USA; 2010.
8. Sakr S, Liu A, Batista DM,Alomari M. A survey of large scale data management approaches in cloud environments.IEEE Communications Surveys and Tutorials.2011;13(3):311–336.
9. Bacardit J, Llor`a X. Large-scale data mining using genetics-based machine learning.Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery; 2013;3(1): 37–61.
10. de la Iglesia B. Evolutionary computation for feature selection in classification problems. Wiley Interdisciplinary Reviews: DataMining and Knowledge Discovery; 2013; 3(6):381–407.
11. Gao BJ, Ester M, Xiong H , Cai JY, Schulte O.The minimum consistent subset cover problem: a minimization view of data mining. IEEE Transactions on Knowledge and Data Engineering,;2013; 25(3): 690-703.
12. Fong S, Deb S, Yang XS, Li J.. Feature selection in life science classification: metaheuristic swarm search. IT Professional; 2014; 16(4): 24-29.
13. Fern´andez A, del R´ıo S, L´opez V.Big data with cloud computing: an insight on the computing environment, MapReduce, and programming frameworks. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery; 2014; 4(5): 380–409.
14. Merelli E, Pettini M, Rasetti M. Topology driven modeling: the IS metaphor.Natural Computing; 2014.
15. Sadeg S, Hamdad L, Benatchba K, Habbas Z. BSO-FS: bee swarm optimization for feature selection in classification. In International Work-Conference on Artificial Neural Networks; 2015; 387-399.
16. Fong S, Wong R, Vasilakos AV. Accelerated PSO swarm search feature selection for data stream mining big data. IEEE transactions on services computing; 2016;9(1): 33-45.
17. Tennant M, Stahl F, Rana O, Gomes JB. Scalable real-time classification of data streams with concept drift. Future Generation Computer Systems; 2017;75: 187-199.
18. K.A.Tupe KA, Wakchaure MA. A Review on Feature Selection Data Stream Mining in Big Data. IJARIIE; 2017;3(1):1581-1584.
19. Sasikala S,Renuka Devi D. A review of traditional and swarm search based feature selection algorithms for handling data stream classification.Third International Conference on Sensing, Signal Processing and Security (ICSSS). IEEE;2017.
20. Gu, Shenkai, Ran Cheng, Yaochu Jin. Feature selection for high-dimensional classification using a competitive swarm optimizer.Soft Computing; 2018; 22(3): 811-822.
21. Manoj R, Joseph, Anto Praveena MD, Vijayakumar K. An ACO–ANN based feature selection algorithm for big data.Cluster Computing; 2018;1-8.

AUTHORS PROFILE

D. Renuka Devi is a research scholar in the Department of Computer Science, IDE, University of Madras, Chennai, India. Her Research interests include Data mining, Machine Learning, Big Data and AI. She has published 4 research articles, including IEEE and Scopus. She has also presented papers at International conferences and received the best paper award.

Dr. Sasikala. S is a Assistant Professor and Research Supervisor in the Department of Computer Science, IDE, University of Madras, Chennai, India. She has been working for more than a decade. Her Research interests include Image, Data mining, Machine Learning, Networks, Big Data and AI. She has published 3 books in the domain of Computer Science and published 19 research articles in leading journals and conference proceedings, including IEEE, Scopus, Elsevier, and Springer. She has also received the best paper awards and Women achievement awards. She currently serves on the editorial board of IIR Journals.