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

Differential Evolution tuned Support Vector Machine for Autistic Spectrum Disorder Diagnosis

Suresh Kumar R, Renugadevi M

Abstract: ***Autistic Spectrum Disorder (ASD) is a brain developmental disorder which weakens the ability to communicate and interact with others. A child with autism spectrum disorder may have different, repetitive patterns of behaviour, interests or activities, including some specific signs. To diagnose the behaviour of ASD and identify the level of disease on the human is still a challenging task for the doctors. Only by the trained and experienced physician can identify the ASD immediately. The data set for autism problem consist of number of causes and the results based on the symptoms for ASD. So, Data mining algorithm is in need to organize and pattern the ASD details. The machine algorithms are available to classify the data in data mining works. In this proposed work, a machine learning algorithm called Support Vector Machine is used to classify the ASD children accurately. SVM is one of the classification algorithms which finding the hyper plane that maximizes the margin between the two classes. Though SVM give better identification of disease, some children have their unique nature which hides their problem of ASD easily. So, to diagnose the problem accurately, the user defined SVM parameters are tuned by optimization algorithm called Differential Evolutionary Algorithm. DE is an optimization algorithm used to find the optimal solution of SVM parameters. Further, to improve the performance of the proposed method, the dimension reduction technique is followed to reduce the SVM and ANN network dimension. The Sequential Feature Selection (SFS) method is applied in this paper, which select the most influenced variables for the output. The reduced network is further classified by ANN and SVM model. The Data set for the ANN and SVM network has been taken from the real records of the multi-specialty hospitals. The SVM and DE optimized SVM results are compared with another classification model called Artificial Neural Networks. The test results show the betterment of DE optimized SVM which give the classification of ASD child very accurately compare with ANN and DE optimized ANN.***

Index Terms: ***ASD, ANN, DE, SVM classification, SFS.***

I. INTRODUCTION

In today's world, the growth of population increases the diseases which cause human death. The diseases are commonly diagnosed by particular tests such as blood test and urine test. There are some other diseases which are not diagnosed by the blood test. The one type of such disease is called Autistic Spectrum Disorder (ASD). Mostly, the ASD symptoms appear at the childhood of first two years and it causes have been linked with genetic and neurological factors [1]. At this age, since the children are not even starting to talk, it is difficult to diagnose the problem accurately. Chakrabarti et al., [2] proposed the first candidate-gene association study of Asperger syndrome (AS) and of autistic traits. In this paper, it is proposed that, the genetic factor is important cause for ASD. 19 genes are significantly associated with Autism Spectrum Quotient in a population sample. Lord C et al., [3] proposed the Autism Diagnostic Observation Schedule-Generic (ADOS-G) MODEL. It consists of four models of 30-minute observational schedule to diagnose the ASD. This proposed model sensitivities and specificities for autism is excellent diagnostic model. C. Johnson et al., [4] proposed the tool kit which contain screening and surveillance tool for ASD children. The screening test results give various actions that has to be followed based on the risk level of ASD. By this literature, it is come to know that, there are various symptoms the doctors need to identify to decide about ASD. But these symptoms are huge and there is some screening test need to do for finding the risk stages of ASD. It's generally based on the level of impairments and how they impact the ability to function [5]. Though the ASD consist of some common signs, the ASD children has unique mixture of symptoms, severity can sometimes be difficult to determine. So, it is necessary to develop on - line tool to diagnose the problem at quick manner.

There are so many on- line classification tool available to classify the disease very fast. C. Amrit et al., [6] proposed the Decision Support System to identify the possible way of child abuse from the child health data of Netherlands. In this paper, authors used data mining technique and machine learning algorithm to detect the patterns of possible child abuse in the given data. FadiThabtah [7] proposed machine learning tools for ASD diagnostic issue. Among all diagnostic algorithm used, SVM is chosen as a best tool for prediction of ASD due to its higher accuracy of classification. In [8], the ASD is classified by three techniques such as ANN, SVM and fuzzy logic. It is proven that the usage of machine learning algorithm is reduced the autism diagnosing time very drastically in [9].

Revised Manuscript Received on 30 July 2019.
* Correspondence Author
Suresh Kumar R*, Department of Computer Science, Sree Saraswathi Thyagaraja College, Pollachi, India.
Dr.Renugadevi M, Department of BCA, Sri Krishna Arts and Science College, Coimbatore, India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC-BY-NC-ND license http://creativecommons.org/licenses/by-nc-nd/4.0/

In this paper the author used ADTree algorithm to limit the available input features. In this paper Autism Diagnostic Observation Schedule (ADOS) module has been developed to classify the data set. And the
classification result of ADOS is compared with [10] and [11]. In [12], it is proposed, combination of different approaches yields with the good features of the methods like genetic, rough sets, bayesian network for the most important items derived from multiple parameters. The two processes involved are attributes reduction and uncertain reasoning technique.

The online techniques have classified the ASD problems quickly, but still the accuracy of the classification is one of the challenging tasks for the Doctors. Some unique character child or the ASD child with some other Disease is not limited by the symptoms considered for autism. At this point, it is not able to classify the autism child accurately only by using this SVM or ANN classifier. So, it is necessary to use some optimization algorithm to tune the SVM parameter to bring the better result. For the authors' knowledge, till now there is no paper related to optimization algorithm based SVM tuning method for the ADS classification problem.

Support Vector Machine is a supervised learning Algorithm applied for both pattern recognition and classification problems [13]. By selecting optimal hyper plane between the vectors, SVM maximize the margins, which give accurate classification results. The non-linear problems also classified based on the kernel transformation in where the vectors are separated in the higher dimension. The optimum separation of the vectors can be done by proper tuning of the SVM parameters such as gamma and sigma value. The Gamma parameter defines how far the influence of margin reaches the vectors and sigma parameter defines how much the misclassification can be avoided [14]. Various classification and extrapolation algorithms are discussed in [15]. So, the proper selection of the parameter gives accurate classification. The proper selection of the SVM parameter can be achieved through optimization algorithm called Differential Evolution Algorithm (DEA).

DEA is a new heuristic Optimization algorithm has few advantages over other optimization algorithm such as fast convergence values and few control parameters [16]. Like Genetic Algorithm it consists of typical selection and crossover process. The main difference between the GA and DEA is, it consists of more than 10 mutation strategy which can give better results based on the application. The DEA tuned SVM can be applied in [17] for energy calculation application. In this paper, the results are compared with ANN and DEA tuned SVM to show the betterment of proposed system. The slurry flow in the pipeline of chemical industry can be identified through the DEA tuned SVM in [18].

The on – line estimation methods give best classification accuracy, but the large number of input variables make the network complex and slow down the processing time. And at the same time, small numbers of input variables are not enough to give good prediction output. So, the input variables should be fit to give best solution. In this proposed work, the Sequential Feature Selection method is applied to identify the influenced variables for the classification. In [19], the medical diabetes classification is done by SFS method. The hand-written digits classification is also done in this paper.

In this proposed method, DEA is used to tune the SVM parameters such as gamma and sigma to give best accuracy in ASD diagnosis. DE optimized SVM network gives accurate classification for ADS. The DEA tuned SVM classification accuracy results are compared with DEA tuned ANN which shows the good accuracy of DEA – SVM.

II. INTRODUCTION TO SVM

Support Vector Machine (SVM) [14] algorithm is used for ASD because of its accuracy in classification for large data even the number of samples are lesser. SVM classifier is predominantly used as for classification algorithms. SVM is a supervised learning algorithm methodology, when the dataset which is used, have both features and class labels. If the features and class labels are not available, they are classified as unsupervised learning. For supervised learning, an SVM classifier builds a model to predict for different cases. If two classes are available, then they are named as binary SVM classifier. Based on its linearity, they are categorised as linear and non-linear classifier. Linear classifier model is used to estimate the straight hyper plane from two classes by maximising the distance from plane to the closest data on the two classes. In non-linear model, dataset is separated into different classes and the data points are plotted in a larger dimension to find the maximum margin.

The main concept of SVM is that when a set of data points are given, which belong to either of two classes, an optimal separating hyper plane for maximizing the distance (from closest points) of either class to the separating hyper plane, and minimizing the risk of misclassifying the training samples and the unseen test samples. The approach of SVM is to formulate a constraint-based optimisation problem, and then to solve it using quadratic programming (QP). Figure 1 gives the diagrammatic representation of the SVM method.
The training data samples near the boundary planes are called as support vectors. Nonlinear data could be mostly dealt by SVM by mapping the species of low dimensionality derived from input to higher dimensionality for effective classification. Even if the classified data are inseparable SVM can work on it. Also, it can work with number of samples and does not depend on the number of genes. This specific characteristic feature of SVM is very much useful for the problems with low ratio of samples to genes. But, as the number of genes is reduced, the SVM performance increases. Mostly Kernel functions are used to solve SVM classification problems.

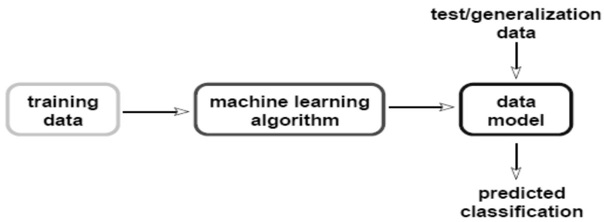


Fig 1. SVM Model

The optimal hyper plane is computed as per the following equation,

$$f(x) = \beta_0 + \beta^\tau \chi \qquad (1)$$

Where, β is known as the weight vector and β_0 is the bias. The hyper plane is given by an infinite way by optimising and . Among various representations, the way chosen for hyperplane is given by equation 2.

$$|\beta_0 + \beta^\tau \chi| = 1 \qquad (2)$$

Where, χ is the shortest distant vector to the hyperplane. Normally these vectors are given as support vectors. This representation is known as the canonical hyper plane. Distance between distant point x and a hyperplane (β, β_0) is given by the equation 3.

$$\text{distance} = \frac{|\beta_0 + \beta^\tau \chi|}{\|\beta\|} \qquad (3)$$

In particular, for the canonical hyper plane, the numerator is equal to one and the distance to the support vectors is

$$\text{Distance}_{\text{support vectors}} = \frac{|\beta_0 + \beta^\tau \chi|}{\|\beta\|} = \frac{1}{\|\beta\|} \qquad (4)$$

The margin is denoted as M, which is twice the distance to the closest data,

$$M = \frac{2}{\|\beta\|} \qquad (5)$$

Finally, the problem of maximizing M is equivalent to the problem of minimizing a function L(β) subject to some constraints. The constraints model the requirement for the hyperplane to classify correctly all the training examples χ_i. Formally,

$$\min_{\beta, \beta_0} L(\beta) = \frac{1}{2}\|\beta\|^2 \text{ subject to } y_i(\beta^\tau \chi_i + \beta_0) \geq 1 \forall i, \qquad (6)$$

where y_i represents each labels of the training examples. This is a problem of Lagrangian optimization that can be solved using Lagrange multipliers to obtain the weight vector and the bias as in Equ. (1) of the optimal hyperplane.

III. DIFFERENTIAL EVOLUTION ALGORITHM (DEA) FOR SUPPORT VECTOR MACHINE (SVM)

Differential Evolution Algorithm (DEA) is an easiest optimisation technique which has a capability to converge quickly for optimum solution. The steps involved in DEA are selection of parameters, mutation, cross over and selection process. The foremost step is to initialise the parameters, which has to be optimised. Those are generated as random variables within limits. After initialisation, fitness values are evaluated for individual population. Next mutation is done by mating the arithmetic combinations of population. It is followed by crossover and selection process. Finally, the best values are selected for next epoch.

The objective of applying DEA is to tune the SVM parameters so as to maximize the classification accuracy. The classification accuracy is determined from the confusion matrix of the testing data set. The representation of confusion matrix is shown in Table 1

Table 1. Confusion matrix.

	Wrongly Predicted	Correctly Predicted
Actual No	True Negative	False Positive
Actual Yes	False Negative	True Positive

The objective function for DE is given by,

$$F(x) = classify(Accuracy) \qquad (7)$$

Where,
Classification Accuracy= (True Positive+ True Negative)/Total no of instances.

In order to effectively analyse the network performance some of the parameters that define the correct classification. They are precision, Recall, F-measures. They are given as,

$$\text{Precision} = \frac{\text{True Positive}}{\text{Predicted } Yes} \qquad (8)$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False } Negative} \qquad (9)$$

$$\text{F-measures} = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}} \qquad (10)$$

The algorithm for implementing DEA for SVM is explained step by step in the following section. The flowchart of the algorithm is also given in fig.2

1. The parameters sigma and gamma are to be optimised. The variables to be optimised are normal integers in DE iteration.

2. For implementing SVM using DEA, accuracy maximization is considered as a fitness function. (Equ. no. 7) This step involves the following steps

SVM method for minimising error is given below.

i. SVM network is initialised by input, training, testing and output data.
ii. Input and output combinations of training data are normalised.
iii. Input and output combinations of testing data are normalised.
iv. Training is done as explained in previous section after the parameters are normalised.
v. Training phase is followed by testing phase where the classification accuracy is calculated.

3. Next step is mutation process wherein randomly selected vector will be mutated.

4. Cross over is done between random vector and mutant vector and a new vector is generated.

5. The selection process selects the vector for surviving for the next generation. It compares the fitness function of the trail vector and target vector.

These steps are continued until prescribed generation is reached. The selected parameters are considered as optimum value when the error is minimised.

The important in machine learning algorithm is feature selection. Feature selection is also called as variable selection or attributes selection. This is mainly done for reduction in the dimensionality to reduce the complexity in computation and to improve the generalisation ability.

Feature extraction and feature selection are the different approaches to reduce the dimensionality.

Feature extraction reduces the dimensionality of the vectors, but the number of features remains same. Feature selection reduces the number of original features but it retains the required information for classification.

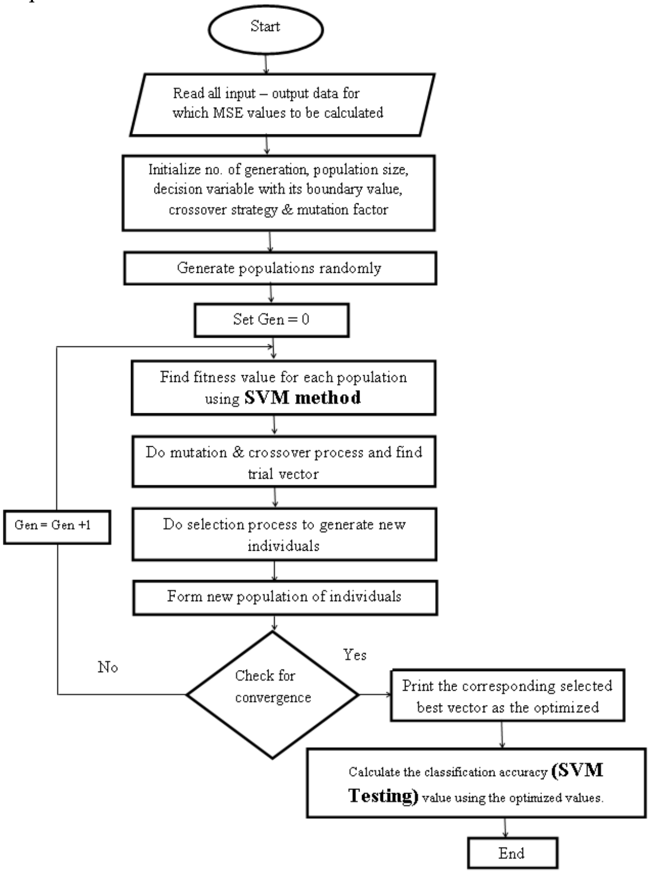


Fig. 2 Flowchart for Proposed DE SVM Algorithm

IV. Feature Selection

For feature selection problem consider a vector whose input is represented as,

$X \in Rk$ and their respective class labels are denoted as $Y \in \{1, -1\}$. Let,

$F=\{f1,f2,...,fk\}$ (11)

This can be taken as the set of all features under test, and let

$S=\{(X(l),Y(l))|l=1,2,...,N\}=\{[x1(l)x2(l)...xk(l)]T,Y(l)|\ l=1, ..., N\}$ (12)

represents the set of training data that contains N pairs, where xi (l) is the numerical value of feature, fi for the ith training sample. The goal of feature selection is to find a minimal set of features

$Fs=\{fs1,fs2,...,fsd\}$ (13)

represent the input vector X in a lower dimensional feature space as

$Xs=[xs1xs2...\ xsd]$ (14)

Where d<k, and the classifier is obtained in the dimension of low range and also the accuracy obtained is improved.

Generally feature selection is categorised into filter methods and wrapper methods. The wrapper method is more efficient than filter method due to the feedback network. Due to its better performance sequential forward feature selection method is used for feature selection.

V. Results and Discussion

Based on methodology proposed in section III, different supervised learning network has been developed for classification of Autism data set. (ASD/No ASD). The Autism screening fortoddlers data set is taken from Kaggle repository. [20]. This data set contains 1054 instances. It has 18 attributes and one class variable. The data class variable is categorised as ASD and no ASD. Hence it has been changed as two variables to represent different class. The class variable ASD has first element as one and second element as zero while no ASD has first element as zero and second element as one.

After reconnoitering a data set it is clear that, the data have 18 attributes and two class variables. The first 10 variables are YES/NO question (1 or 0), next 8 attributes are age, Score by Q-chat-10, Sex, Ethnicity, born with jaundice, Family member with ASD history, who is completing the test and Why are you taken the screening respectively. All this variable is taken as input for developing the network. Out of 1054 instances, 73 % data is used for training and the remaining 27 % data is used for testing. The simulation is carried for entire data set as well as reduced data set. Sequential Forward feature Selection method is used in proposed work and the results are compared with the Random forest method. The network is developed by using MATLAB.

There are totally eight different networks have been developed for classifying Autism data set. They are,

Network 1: Artificial neural network (ANN) without feature reduction.

Network 2: Support Vector Reduction (SVM) without feature reduction.

Network 3: DE tuned ANN without feature reduction.

Network 4: DE tuned SVM without feature reduction.

Network 5: ANN with feature reduction.

Network 6: SVM with feature reduction.

Network 7: DE tuned ANN with feature reduction.

Network 8: DE tuned SVM with feature reduction.

The performance of all this developed network are analysed by classification accuracy, Precision, Recall, F-Measures and simulation time.

Simulation Study

This following discusses the performance of different developed network for classifying autism data set.

A. Artificial neural network (ANN) without feature reduction.

In this section the performance of ANN without data reduction is presented. The network has 17 inputs and two class variables (ASD, no ASD). The network is trained with scaled conjugate gradient method up to the stopping criteria. Entire 1054 instances are used for network development in which 778 (73 %) used for training and 276 (27 %) used for testing (30%) the network.

The parameters used for ANN development and testing performance of developed ANN network is given in Table.2

Table 2: Performance of ANN for Autism data set classification

Network Parameters			Performance Analysis				
Number of Hidden layer	Learning Rate	Momentum factor	Accuracy	Precision	Recall	F-measures	Training time in sec
6	0.5	0.3	87.32 %	84.93%	72.09%	77.09 %	48.5

The Network parameters such as number of hidden layer, learning rate and momentum factor are user defined variables. The values for the variables are chosen randomly. From this performance analysis it is inferred that ANN classify the Autism data not up to the level. The training time taken for the classification is also 48.5 sec which is long for big data set.

The Regression plot and confusion matrix of ANN classifier is shown in fig 3. The regression value is also 0.808 that shows average performance of ANN. The regression plot gives how close the output values are with target values. From this figure it is inferred that how much output value outlier with target value. The confusion matrix shows the performance of the ANN model for the given classifier set.

The confusion matrix analyses the test data set and gives the detail about what is correct in the performed model and what error the network has been done. From the matrix it is concluded that the right work done in the model is only 87.3%

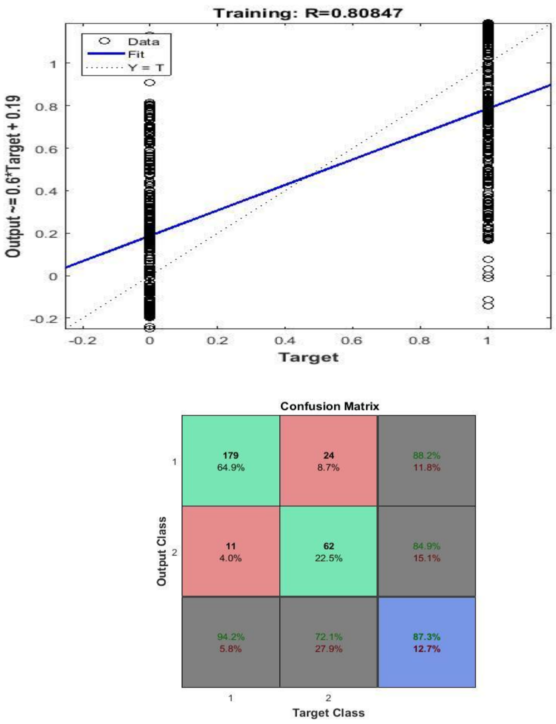


Fig.3 Regression and Confusion matrix

B. Support Vector Reduction (SVM) without feature reduction.

This section provides the performance analysis of Support Vector classifier for Autism data set. The same data set is used for SVM classifier also. The network is trained with algorithm presented in section III.

The network parameters and performance of SVM classifier is given in Table 3. The network parameters gamma and sigma are chosen randomly. Radial basis function classifier used in SVM training.

Table 3: Performance of SVM for Autism data set classification

Network Parameters		Performance Analysis				
Gamma	Sigma	Accuracy	Precision	Recall	F-measures	Training time in sec
1	4	94.57%	100%	82.56 %	90.45%	2.3

The Accuracy and precision are far better than ANN but it not only guaranteed the best performance because the remaining performance parameters such as Recall and F-measures are not upto the level. The Regression plot, Confusion matrix and RoC are shown in fig 4. The Regression value of SVM network is increased slightly compared with ANN network. The confusion matrix gives 94.6% of correct classification which is more than ANN network.

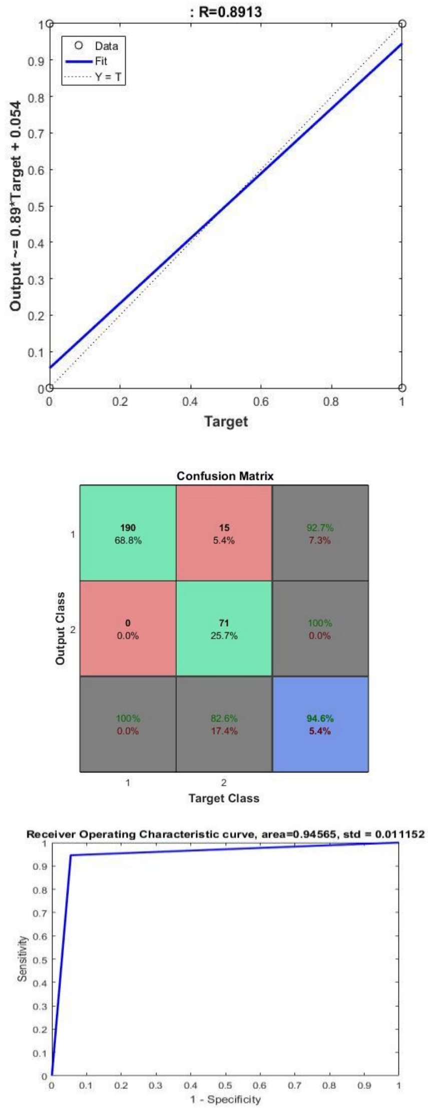


Fig 4. Regression plot, Confusion matrix and RoC of SVM Classifier.

From this fig 4, it is clear that the regression is only about 0.89 it also indicates the classifier is performing better.

C. Network 3: DE tuned ANN without feature reduction.

One of the challenges in ANN classifier is selection of network parameters such as number of hidden layers, learning rate and momentum factor. All these parameters are influence the performance of ANN. Proper selection of all these parameters improve the network performance. In this proposed work, Differential Evolution (DE) algorithm is adopted to find the optimal value of ANN network parameters. The DE parameters are given below.

Table 4: Parameters of DEA

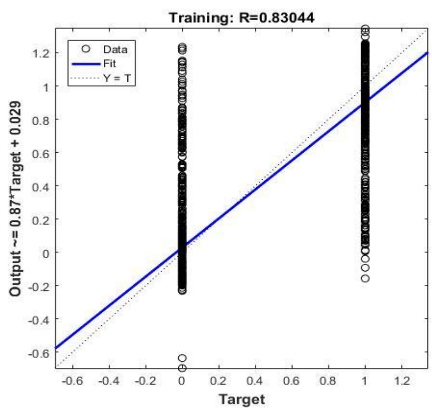
Parameters of DE	
Number of populations	50
Maximum number of iterations	100
Mutation Strategy	DE/rand/bin/1
Scaling Factor F	0.5
Crossover rate CR	0.8
Number of Decision variable	2

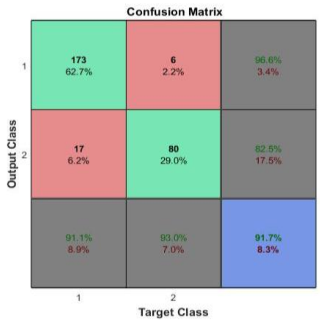
The table 5 shows the optimal network parameters obtained by DE and performance of DE tuned ANN classifier for Autism data set. The same data set is used in this network also. The upper and lower limit for the number of hidden layer, learning rate and momentum factor are 1-100, 0-1 and 0-1 respectively.

Table 5: Performance of DE tuned ANN for Autism data set classification

Optimal Network Parameters Obtained from DE			Performance Analysis				
Number of Hidden layer	Learning Rate	Momentum factor	Accuracy	Precision	Recall	F-measures	Training time in sec
16	0.5383	0.9102	91.67%	82.47%	93.02%	87.43%	210

The accuracy of DE tuned ANN is increased with optimal network parameters. The remaining parameters such as Precision, Recall and F-Measures also improved when compared to normal DE tuned ANN classifier. But again, accuracy is only about 92 %. So, it is necessary to develop a network for more accurate classification. But the training time is high because DE is iterative method. It takes time to converge. This is the drawback in tuning of network parameters. The Regression plot, Confusion matrix and RoC for DE-ANN classifier is shown in fig. 5. From this figure, it is evident that the regression level still needs to be improved.





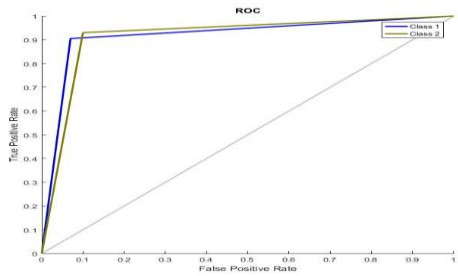


Fig. 5 Regression plot, Confusion Matrix and RoCfor DE-ANN classifier

D. DE tuned SVM without feature reduction.

In this section, the performance of DE-SVM classifiers for Autism data set is presented. In SVM also the user defined parameters such as gamma and sigma have to be carefully chosen. These two parameters have great impact on SVM performance. In this work, the DE optimization is applied to find the optimal parameters of SVM in order to improve the accuracy of the classifier.The optimal parameters of SVM and the testing performance of DE SVM classifier is given in Table 6. The entire Autism data set is used for network classifier.

Table 6: Performance of DE-SVM for Autism data set classification

Network Parameters		Performance Analysis				
Gamma	Sigma	Accuracy	Precision	Recall	F-measures	Training time in sec
1	8.617986	98.9130 %	100%	96.51 %	98.22 %	128

The network performance improved noticeably, when SVM parameters is tuned with DE. Table 6 shows the improved values of accuracy and precision. The accuracy and all other parameters are reasonably good for classification of Autism data set. Training time is slightly higher because of DE iterative process.

The regression plot and confusion matrix are shown in fig 6. From this figure it is clear that regression also improved well with DE based SVM classifier. The convergence plot of DE-SVM is also given in fig 7. This graph shows the good convergence behaviour of DE-SVM.

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

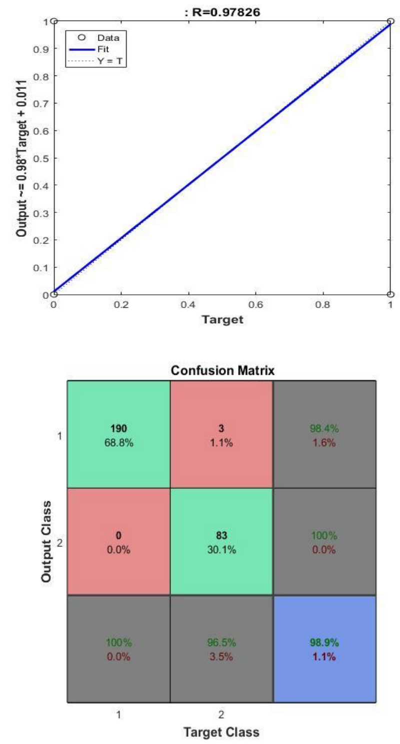


Fig 6. Regression and Confusion matrix for DE SVM Classifier

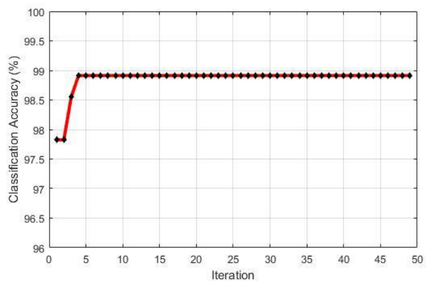


Fig 7. Convergence curve of DE-SVM classifier

E. Feature Selection for Autism data set

When developing the machine learning based classifier network, data plays major role. The classifier performance and complexity depend on the size of training data set. Larger data set makes networks learns accurately. But it makes network more complex and also it increases training time. Hence the real time implementation of machine learning classifier is more complex.

The reduction of data set makes the system simple but at the meantime reduction of data set cannot affect the classifier performance. Hence, it is essential to identify the impactable attributes for the class variable. Only these impactable variables used in network development. In this work, Forward Sequential search Method is adopted in order to extract the features. Based on the method proposed in section IV, the features for Autism data set is extracted from the input. Totally 1054 instances with 18 attributes with 2 class variables are given as input to Sequential Feature Selection method.

The selected features from the Sequential Feature Selection method are shown in fig. 8. Out of 18 features only 3 features are selected. The selected features are 3, 10 and 12. The selected features are listed in Table 7.

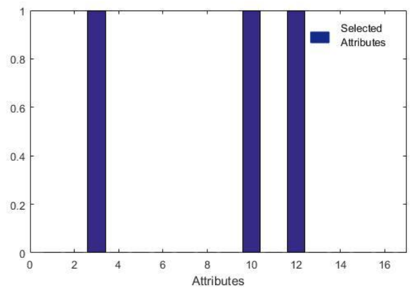


Fig 8. SFS Selected Features

Table 7. Sequential Feature Selection Performance for Autism data set

No. of components in PATTERN vector	No. of components in FEATURE vector	Dimensionality Reduction	Execution Time for SFS Method
18	3 (3,10,12)	82.3529%	1.0039 secs

From this result it is inferred that the Dimensionality of the network almost reduced to 82%. It makes the data set very simple. Now, the data set has only 3 attributes. The performance of this method is tested by both ANN and SVM and its performance is given in next section.

F. ANN with feature selection

In order to reduce the network complexity and improve the performance of the classifier, the data set has been reduced with feature selection method. The selected feature (3, 10, and 12) is only applied to train the classifier. In this section the performance of ANN classifier with reduced data set is presented. The ANN classifier used, defined parameters such as number of hidden layers, learning rate and momentum factors are chosen randomly.

The network parameters and performance of the classifier is given in Table 8. From this table it is clear that network classify the data accurately even with reduced data. The training time also reduced compared to ANN classifier with full data set. The regression plot and Confusion matrix for the ANN classifier is shown in fig.9. From this graph it is evident that regression also improved compared with normal ANN classifier.

Table 8: Performance of ANN for Autism data set classification with feature selection

Network Parameters			Performance Analysis				
Number of Hidden layer	Learning Rate	Momentum factor	Accuracy	Precision	Recall	F-measures	Training time in sec
6	0.5	0.3	94.20 %	100 %	81.40 %	89.74 %	21

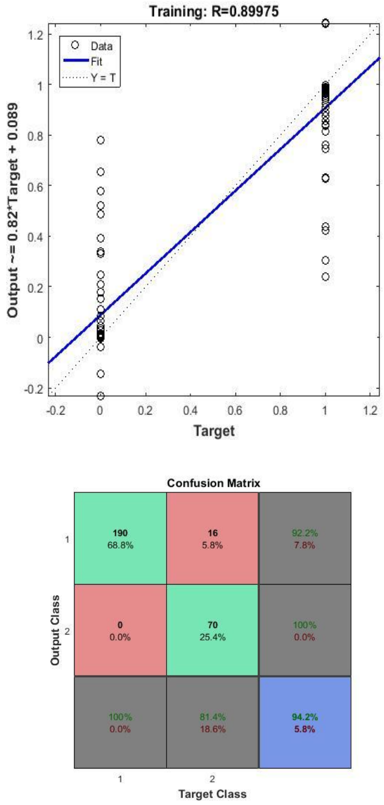


Fig. 9 Regression Plot and Confusion matrix of ANN with feature selection

G. SVM with feature selection

In this section the performance of SVM with feature selected data is presented. The SVM classifier is trained with 3 selected features and two class variables. The user defined data for SVM such as gamma and sigma are chosen randomly.

Table 9 shows the randomly chosen user defined parameters and its performance analysis. From this table it is inferred that the performance improved better compared to normal SVM. The regression plot and confusion matrix given in fig 10. This is also indicating the better performance of SVM classifier.

Table 9: Performance of SVM for Autism data set classification with feature selection

Network Parameters		Performance Analysis				
Gamma	Sigma	Accuracy	Precision	Recall	F-measures	Training time in sec
1	0.01	99.64%	100%	98.84%	99.42%	2

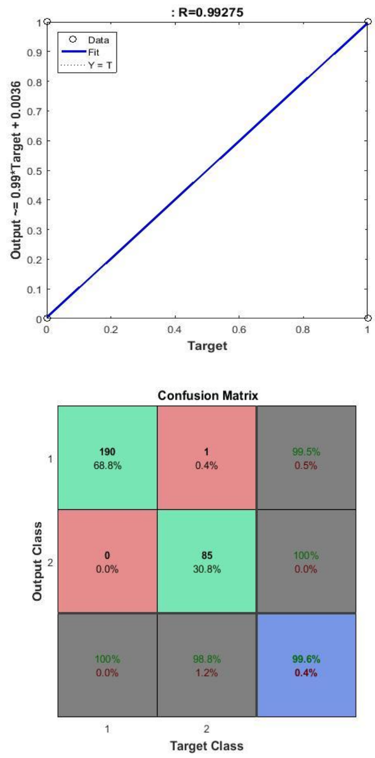


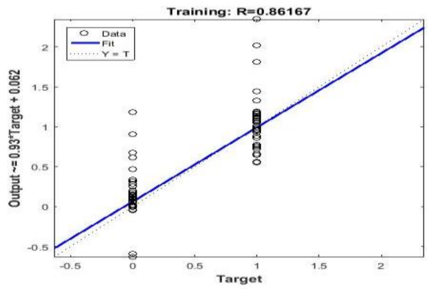
Fig 10. Regression plot and confusion matrix of SVM with feature selection

H. DE tuned ANN with feature reduction

In order to reduce the difficulty in choosing the user defined parameters for ANN, DE is applied to tune ANN parameters. The ANN classifier is trained with selected features. The table 10 shows the optimal parameters and performance analysis of DE tuned ANN classifier. From this table, it is clear that the network performance is better. The figure 11 shows the Regression plot and confusion matrix for DE-SVM classifier with feature selection.

Table 10: Performance of DE tuned ANN for Autism data set classification with feature selection

Optimal Network Parameters Obtained from DE			Performance Analysis				
Number of Hidden layer	Learning Rate	Momentum factor	Accuracy	Precision	Recall	F-measures	Training time in sec
36	0.7687	0.0595	97.46%	100%	91.86 %	95.76 %	180



Retrieval Number: B3063078219/19©BEIESP
DOI: 10.35940/ijrte.B3063.078219
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

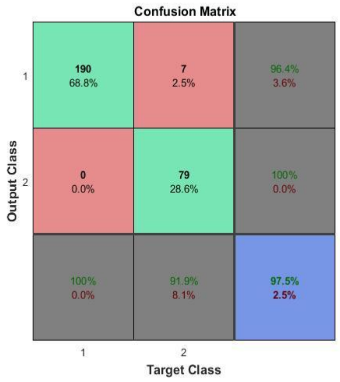


Fig 11. Regression plot and confusion matrix for DE-ANN classifier with feature selection

I. DE tuned SVM with feature reduction

Finally, the DE tuned SVM classifier is developed for more accurate classification of reduced data set. In this work, SVM parameters such as gamma and sigma are tuned with DE. The SVM trained with selected features. Hence network has 1054 instances with 3 inputs and two class output.

The performance and optimal parameters of DE tuned SVM with feature selection is given in table 11. From this table it is clear that, all network performance parameters are 100%, which means this network is superior classifier network for Autism data set. The fig 12 shows the regression and confusion plot of DE SVM with feature selection. From this graph, it is evident that the regression value is one, which means, network works very accurately.

Table 11: Performance of DE-SVM for Autism data set classification

Network Parameters		Performance Analysis				
Gamma	Sigma	Accuracy	Precision	Recall	F-measures	Training time in sec
23.636185	9.013184	100%	100%	100%	100%	105

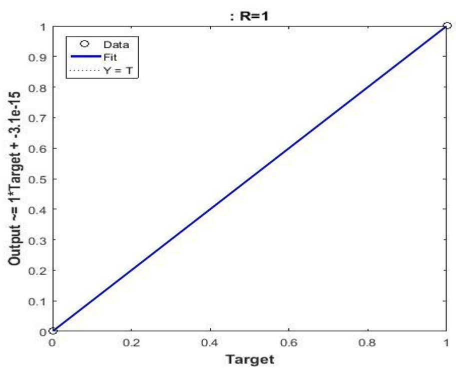


Fig 12. Regression plot and confusion matrix for DE-SVM classifier with feature selection

VI. CONCLUSION

In this paper, Machine learning tool is used for Data mining application. The Autistic Spectrum Disorder (ASD) data set has large dimension matrix. The ASD data set is applied in SVM network to find the ASD classification. The accuracy, precision, recall, F measures and Timings of the network have been analysed. The user defined ANN and SVM parameters are optimized through Differential Evolutionary Algorithm. DE tuned SVM network has given good performance compared with simple ANN network, simple SVM network and DE tuned ANN network.The ASD data set is further reduced by SFS feature selection method. Based on the impact of input variables on the output variable, 18 input attributes of ASD network is reduced to 3. Hence, for the reduced network, the DE tuned SVM model is applied and the performance analysis have been made. The results show the better performance of proposed DE tuned SVM network with future reduction model compared with reduced feature of simple ANN, Simple SVM and DE tuned ANN network. This Online ASD classifier can be used in the real time diagnosing application. In future, the usage of Online ASD classifier will reduce the ignorance of the disease and give the initial awareness about the disease. This proposed method is helpful for the physicians to identify the disease very soon with more accurately. Once the Autism is identified at earlier stage with perfect level, it is better for physician to give the medicine accordingly, which will reduce the number of ASD patient in India.

REFERENCES

1. Bolton P, Macdonald H, Pickles A, Rios P, Goode S, Crowson M, Bailey A, Rutter M. A case-control family history study of autism. J Psychol Psychiatry. 1994;35(35):877–900. doi: 10.1111/jcpp.1994.35.issue-5.
2. Chakrabarti, B, Dudbridge, F, Kent, L, Wheelwright S, Hill-Cawthorne G, Allison C, Banerjee-Basu S, Baron-Cohen S, et al. Genes related to sex steroids, neural growth, and social emotional behavior are associated with autistic traits, empathy, and Asperger syndrome. Autism Res. 2009;2(3):157–77. doi:10.1002/aur.80

3. Lord C, Risi S, Lambrecht L, Cook EH Jr, Lambrecht BL, DiLavore PC, Pickles A, Rutter M. et al. the autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. J Autism DevDisord. 2000;30(30):205–23. doi:10.1023/A:1005592401947.
4. C. Johnson, S. Myers, P. Lipkin, J. Cartwright, L. Desch, J. Duby, E. Elias, E. Levey, G. Liptak, N. Murphy, A. Tilton, D. Lollar, M. Macias, M. McPherson, D. Olson, B. Strickland, S. Skipper, J. Ackermann, M. Del Monte, T. Challman, S. Hyman, S. Levy, S. Spooner, and M. Yeargin-Allsopp. Identification and evaluation of children with autism spectrum disorders. Pediatrics, 120(5):1183–1215, 2007.
5. "Autism spectrum disorder" https://www.mayoclinic.org/
6. C. Amrit, T. Paauw, R. Aly, and M. Lavric. Using text mining and machine learning for detection of child abuse. CoRR, abs/1611.0:31, nov 2016.
7. FadiThabtah, "Machine learning in autistic spectrum disorder behavioral research: A review and ways forward" INFORMATICS FOR HEALTH & SOCIAL CARE 2017, 1–20., Taylor and Francis
8. Mythili M, Shanavas Mohamed R. A study on Autism spectrum disorders using classification techniques. Ijcsit. 2014;5(6):7288–91
9. Bone D, Goodwin MS, Black MP, Lee -C-C, Audhkhasi K, Narayanan S. "Applying machine learning to facilitate autism diagnostics: pitfalls and promises". J Autism DevDisord. 2014;45(5):1–16.
10. Wall DP, Kosmicki JA, DeLuca T, Harstad EB, Fusaro VA. Use of Machine Learning to Shorten Observation-based Screening and Diagnosis of Autism. Translational Psychiatry. 2012a;2(4):e100.
11. Wall DP, Dally R, Luyster R, Jung JY, DeLuca TF. Use of Artificial Intelligence to Shorten the Behavioral Diagnosis of Autism. PloS one. 2012b;7(8)
12. Yin Z, Zhao Y, Lu X, Duan H, "A hybrid intelligent diagnosis approach for quick screening of Alzheimer's disease based on multiple neuropsychological rating scales" Comput Math Methods Med. 2015;2015:258761. doi: 10.1155/2015/258761. Epub 2015 Mar 1.
13. Vapnik V. The Nature of Statistical Learning Theory. New York: Springer Verlag; 1995.
14. Vapnik V. Statistical Learning Theory. New York: John Wiley; 1998.
15. Dr.A. Sheela, A.Madhumitha, Dr.S.Vijayachitra," Comparison of extrapolation techniques for active and reactive power prediction in grid connected inverter" International Journal of Computer Informatics & Technological Engineering, volume 2 issue 1, January 2015
16. Storn R, Price K. Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. Journal of Global Optimization. 1997; 11:341–359.
17. AIS Velusamy, NB Ramu, D Durairaj, K Murugesan, "Differential evolutionary algorithm-based optimal support vector machine for online dynamic available transfer capability estimation incorporating transmission capacity margins" International Transactions on Electrical Energy Systems 27 (7), e2331
18. Lahiri SK, Ghanta KC. The support vector regression with the parameter tuning assisted by a differential evolution technique: study of the critical velocity of a slurry flow in a pipeline. Chemical Industry and Chemical Engineering Quarterly. 2008; 14(3):191–203.
19. Thomas R¨uckstieß, Christian Osendorfer, and Patrick van der Smagt, "Sequential Feature Selection for Classification" Australasian Joint Conference on Artificial IntelligenceAI 2011: Advances in Artificial Intelligence pp 132-141
20. Fadi Fayez Thabtah Department of Digital Technology, Auckland, New Zealand, https://www.kaggle.com/fabdelja/autism-screening-for-toddlers

AUTHORS PROFILE

Mr.R.Suresh Kumar is working as Assistant .Professor 14 years, Department of Computer Science, Sree Saraswathi Thyagaraja College. His research interests include classification, clustering of Data Mining for different type of datasets.

Mrs. Dr.M.Renuka Devi , has nearly 18 years of post graduate teaching experience in Computer Science. She has indulged in training the post graduate students to complete real time projects and also guides research scholars in Computer Science. She has published 40 papers in various international journal. There are 8 scholars doing PhD under her guidance. She has acted as chair in international conferences. Currently she is working as Professor and Head in the Department of BCA at Sri Krishna Arts and Science College.