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

Optimization of Machine Learning Techniques on Autism Spectrum Disorder with Swarm Intelligence Based Feature Selection

K. Vijayalakshmi, M. Vinayakamurthy, V. Anuradha

***Abstract:** This paper is a study on the various machine learning algorithms in order to perform ASD (Autism spectrum Disorder) as per the DSM-V standards. ASD occurs more frequently among children and in order to diagnose this with better accuracy, the study on binary firefly algorithm, a swarm intelligence based wrapper feature selection algorithm is used to obtain best results with optimum feature subsets. This paper will provide overall result after applying it to all types of machine learning models on supervised learning.*

***Index Terms:** ASD, Binary firefly, Machine learning, Swarm Intelligence..*

I. INTRODUCTION

There is a sustainable growth in the victims that are diagnosed with autism disorder in the recent decade. Autism is very common among younger generations [1]. It is a childhood disorder. According to the survey in USA, it is observed that 1 among every 86 children within the age of 8 is suffering or diagnosed with this disease. According to the DSM-V Standards[2], an autism diagnosis is conducted which is called as clinical examination procedure. This procedure helps in classifying the disorder. As per the contributions and experiences, US Mental Health professionals formulated these standards that are widely used for classifying ASD from non-ASD in the behavioral analytics. These standards involves questionnaire and interview which are carried out for classification of behavior. There are two types of common tests which are performed by pedestrians in order to detect the symptoms of childhood autism. These tests are ADOS and ADI-R, which are performed by certified professionals under controlled environment such as laboratories.

Revised Manuscript Received on 30 July 2019.
***** Correspondence Author
K. Vijayalakshmi*, M.C.A degree from University of Bharadhidasan, Trichy India,
Dr. M. Vinayakamurthy, Professor in School of Computer Science and Applications, REVA Unviersity, Bangalore, Karnataka, India.
Dr. Anuradha, HOD, PG in a reputed institution, Tamil nadu, India.

These tests last 60 minutes of duration that also relies over the responsiveness of patients. On the basis of response quality, the binary scores are assigned by these professionals which are further analysed to know the autism severity in the patients[3].

In ASD dataset, 21 behavioral attributes is taken for classification.

This technology also advanced the screening approach over handheld devices such as mobile phones by mobile application based ASD which is also designed with respect to DSM-V standards in terms of autism detection[4]. The procedure of acquiring the better accuracy of subset along with optimum feature is known as dimensionality reduction[5]. There are two approaches of dimensionality reduction out of which feature selection is recommended as it is best suited for real world datasets[6]. ASD tasks for Autism Spectrum Disorder dataset involves 21 features which makes it highly dimensional dataset. It is stated that even 5 attributes are sufficient for efficient classification [7]. The behavior patterns can be formed by classifying the differences of ASD patient from ADHD[8]. This paper will include the adequate details about the work carried with Pre-processing tasks and datasets. The future scope is also elaborated within the paper after interpreting the results.

II. PROPOSED WORK

In ASD dataset, there exist 1 class attribute and 20 features within 21 attributes. This dataset can generate 2 to the power 20 feature subsets for evaluation purpose. In search based selection algorithms such as exhaustive search will face exponential increase the time complexity whereas the sub-selection is classified as np-hard problem which can be overcome by feature elimination algorithms with evaluation of candidate. This is the best solution for np-hard problem. A feature selection strategy is formulated by integrating the feature elimination approach and selection based approach as per the ranking. But this ranking approach caused redundancy among features which affected the machine learning performance as it is highly prone to inter-feature correlation. Therefore, in order to have best alternative swarm intelligence based feature selection wrappers are proposed in order to perform the feature evaluation function. Swarm intelligence wrappers are considered as better feature selection explorers as it provides more opportunity to explore many possibilities with maximum result and within minimum iterations which is the major objective of selection [9].

In order to make it more effective single objective function of firefly algorithm is combined with swarm algorithm so that maximum accuracy can be achieved with minimum fitness of subsets features [10].

A. Feature Selection

Technically, feature selection is a process of selecting a candidate subset of features or relevant features [11]. This criteria is used in order to obtain optimal feature subset. There are three main techniques of feature selection which are as follows such as embedded methods, wrappers and filters. For ASD, wrappers is more suitable as it is also called as evaluation method. In first module, results are derived by using 8 different algorithms of machine learning, the ASD children dataset is trained by using fold cross validations around 10 and are compared with the machine learning model results obtained by feature selection algorithm which is deployed with selection algorithm of binary firefly feature. The pseudo code and configuration is mentioned in the following. For optimization, the binary firefly algorithm is implemented and developed as feature selection algorithm.

Wrapper Based Feature Selection Binary firefly selection wrapper is introduced to optimize the regression and classification algorithms [12]. This algorithm has been proved to be a fast performer and outperformed benchmark algorithms over 40 datasets. Using logistic chaotic map [13], the binary firefly algorithm is boasted to increase its attractiveness. By simulated annealing, the global as well as local search strategy of feature selection is enhanced. Therefore, this algorithm suites best for ASD solution as it includes minimum iterations in overall coverage. This algorithm is considered as a swarm intelligence optimizer [14] as it involves single objective function. Following is the general architecture of the work in Fig.1. In this, the architecture of the proposed work is provided to evaluate feature subset along with single objective over the ASD datasets. Also it shows clearly the how firefly swarm intelligence based feature selection algorithm is incorporated into the proposed work to obtain the optimum features from the given dataset to maximize the accuracy of classifiers.

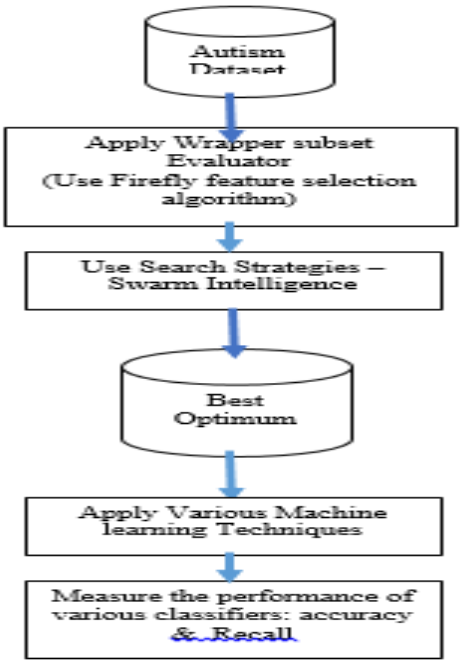

Fig 1 – Framework of the proposed work

III. PRE-PROCESSING WITH SUPERVISED MACHINE LEARNING TECHNIQUES

The parameters considered for wrapper feature selection and Machine learning is provided in Table 1. 10 featured subsets are selected using binary firefly feature selection, out of 21 features as optimum dataset.

Apply Firefly wrapper swarm intelligence based feature selection algorithm using mining tools to optimize the features of a given dataset, global solutions, max iterations, best solutions, number of subsets,chaotic &objective function

Table 1 :Swarm Intelligence based Feature selection parameters

The various classifier based machine learning algorithms are:

a. J48 as a decision based classifier
b. Naïve Bayes as probability based classifier
c. K-NN as distance based classifier
d. SVM as discriminate function based classifier

These algorithms are used for the derived optimum dataset and measured the performance metrics according to the objective of this work.

a. Accuracy : % of correctly classified records

b. TP rate/Recall: predicted positive instances correctly: predicted negative instances.

Table 2: Performance Metrics

In Table 2, the parameters to analyse the performance of various algorithms of machine learning on the ASD children dataset is mentioned.

IV. RESULTS

In Table 3, the results of those parameters in table 2 is listed by testing the dataset before and after applying Binary Firefly algorithm of feature selection.

	Recall		Accuracy	
	Before	After	Before	After
J48	0.91	0.92	91	92
NB	0.93	0.96	93	96
KNN	0.88	0.93	88	94
SVM	1.00	0.98	99	98

Table 3: Performance of Machine learning classifiers

In table 3, it is observed that on original dataset the range of accuracy of various machine learning models on ASD is nearly from 88% to 99% before applying swarm intelligence based feature selection algorithm.

The least accuracy is produced by K-NN classifier with 88% and maximum accuracy by SVM.

Naive Bayes classifier [15] and decision tree provided medium accuracy of performance. Minimum scores is achieved by SVM and MLP classifiers are 0.06 and 0.05 respectively with maximum ROC of 1 [16]. All these algorithms are unsuccessful in terms of achieving the true positive rate whereas other algorithms have misclassification errors. The overall accuracy obtained from feature selection features which are decreased to 10 are in the range of 92%-98%. K-NN model produced more accurate results with improvement of 6.% than the trained model performed over the original dataset. Hence, in overall scenario we can see that all the three classifiers have shown improvement in accuracy with optimum results than SVM models.

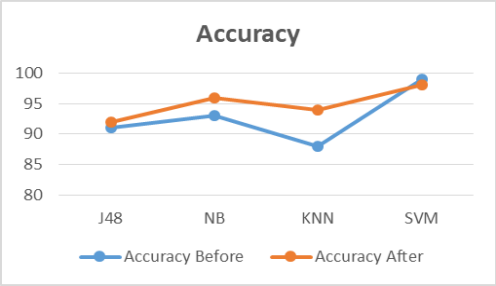

Fig 2 – Accuracy

In Fig. 2, the graphical view of variation in terms of accuracy after and before binary firefly feature selection within machine learning models. From the figure it is clear that the after feature selection that models are able to perform better or perform closer to the machine learning models built on entire datasets.

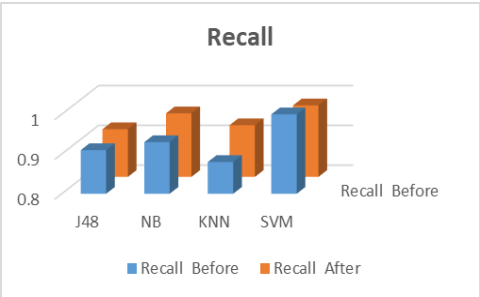

Fig 3 – Recall

In Fig. 3, the recall /TP Rate obtained by the various machine learning models shown with optimization. This method is given more importance as false classification of positive instances impact are false. This can cause the harmful effect over the prediction. For example, if this classification method classifies that kid is not suffering of autism whereas he is a victim of this disease then it can delays diagnostic process and affect the treatment process of that kid which can increase the complications. Therefore, in case of medical datasets, recall rate is considered as the most favorable evaluation factor as it can provide more accurate results than any other factors. The classification of false positive rates can be negotiated as it can be proved wrong by the therapist's after the diagnosis. This cannot disturb the process of diagnosis of the child. It is observed that there is an improvement in this measure after feature selection, but the j48 and SVM classifiers are considerably closer.

V. DISCUSSIONS

After the evaluation, it is observed that 55% - 45% of records classified between yes and no class which clearly shows that both classes are equally influenced by the characteristics of Autism dataset. Stochastic Swarm intelligence algorithms is the best choice as it has more search capability with maximum number of iterations. Binary Firefly algorithm, in the other hand, is the fast exploring algorithm for feature selection. The amount of dataset instances are less which provides a chance to model to over fit the dataset. The interpretation had stated that prediction performance of the machine learning models has improved by optimum behavior set with a minimum features and decent performance.

VI. CONCLUSIONS

The objective of the paper work is to give confidence on the reduced feature set algorithms performance on various machine learning algorithms. This model shows that there can be an improvement in the prediction due to the less number of behavior sets of Autism which in deed help for better diagnosis. In future work, the framework can be implemented for better preprocessing to make the datasets suits for all types of supervised and unsupervised machine learning techniques.

REFERENCES

1. Underwood, E. (2016). People on autism spectrum die 18 years younger than average. Science.
2. Grzadzinski, R., Huerta, M. And Lord, C. (2013). DSM-5 and autism spectrum disorders (asds): an opportunity for identifying ASD subtypes. Molecular Autism, 4(1), p.12.
3. Levy, S., Duda, M., Haber, N. And Wall, D. (2017). Sparsifying machine learning models identify stable subsets of predictive features for behavioral detection of autism. Molecular Autism, 8(1).
4. Harstad, E., Fogler, J., Sideridis, G., Weas, S., Mauras, C. And Barbaresi, W. (2014). Comparing Diagnostic Outcomes of Autism Spectrum Disorder Using DSM-IV-TR and DSM-5 Criteria. Journal of Autism and Developmental Disorders, 45(5), pp.1437-1450.
5. Saxena, A. And Wang, J. (2010). Dimensionality Reduction with Unsupervised Feature Selection and Applying Non-Euclidean Norms for Classification Accuracy. International Journal of Data Warehousing and Mining, 6(2), pp.22-40.
6. Kosmicki, J., Sochat, V., Duda, M. And Wall, D. (2015). Searching for a minimal set of behaviors for autism detection through feature selection-based machine learning. Translational Psychiatry, 5(2), pp.e514-e514.
7. Jaba, S. And Shanthi, V. (2009). An Approach for Discretization and Feature Selection Of Continuous-Valued Attributes in Medical Images for Classification Learning. International Journal of Computer and Electrical Engineering, pp.179-183.
8. Knopf, A. (2019). ADHD/ASD in families: Sibling study shows differences and similarities of disorders. The Brown University Child and Adolescent Behavior Letter, 35(2), pp.3-3.
9. Senthilkumar, S. (2014). Practical Applications of Swarm Intelligence and Evolutionary Computation. International Journal of Swarm Intelligence and Evolutionary Computation, 03(02).
10. Revathi, A. And Sumathi, D. (2017). Finding Accuracy in Feature Selection Using Firefly Algorithm with Rough Set theory. International Journal of Innovative Research in Computer Science & Technology, 5(6), pp.398-402.
11. Feature selection method for High Dimensional Data. (2016). International Journal of Modern Trends in Engineering & Research, 3(10), pp.190-203.

12. Liu, L. And Miao, S. (2015). The complexity of binary sequences using logistic chaotic maps. Complexity, 21(6), pp.121-129.
13. Torgo, L. And Gama, J. (1997). Regression Using Classification Algorithms. Intelligent Data Analysis, 1(4), pp.275-292.
14. Brezočnik, L., Fister, I. And Podgorelec, V. (2018). Swarm Intelligence Algorithms for Feature Selection: A Review. Applied Sciences, 8(9), p.1521.
15. Inza, I., Larrañaga, P., Etxeberria, R. And Sierra, B. (2000). Feature Subset Selection by Bayesian network-based optimization. Artificial Intelligence, 123(1-2), pp.157-184.
16. Radhamani, E. And Krishnaveni, K. (2016). Diagnosis and Evaluation of ADHD using MLP and SVM Classifiers. Indian Journal of Science and Technology, 9(19).
17. Erskine, A., Joyce, T. And Herrmann, J. (2017). Stochastic stability of particle swarm optimisation. Swarm Intelligence, 11(3-4), pp.295-315.
18. Pratap, A. And Kanimozhiselvi, C. (2014). Predictive assessment of autism using unsupervised machine learning models. International Journal of Advanced Intelligence Paradigms, 6(2), p.113.
19. Hull, R. (2012). Ranking algorithms. New Scientist, 215(2881), p.28.
20. Yang, X. (2014). Diversity and Mechanisms in Swarm Intelligence. International Journal of Swarm Intelligence Research, 5(2), pp.1-12.

AUTHORS PROFILE

K. Vijayalakshmi received M.C.A degree from University of Bharadhidasan, Trichy India, in 1999, and the M.Phil degree in Computer Science from Manonmaniam Sundranar University, Trinelveli, India, in 2002. Currently she is pursuing Ph.D programme in Computer Science and Engineering at REVA University, Bangalore, Karnataka, India. She has presented and published research papers in national conferences & International journals. She is a life member of ISTE (Indian Society for Technical Education)

Dr. M. Vinayakamurthy is a Professor in School of Computer Science and Applications, REVA Unviersity, Bangalore, Karnataka, India. He has more than 30 years of teaching and research experience. His areas of specialization include Data mining, Fuzzy logic and Optimization Techniques. He has over 50 publications in International Conferences and Journals to his credit. He has also published books in the field of Mathematics.

Dr. Anuradha is currently working as HOD, PG in a reputed institution, Tamilnadu, INDIA. She has 20 years of experience in teaching with research experience. She published and presented in various conferences and journals both at national and international.