Classification of Autistic Spectrum Disorder Using Deep Neural Network With Particle Swarm Optimization

Sanat Kumar Sahu, Govt. K. P. G. College, Jagdalpur, India
Pratibha Verma, Dr. C. V. Raman University, Bilaspur, India

ABSTRACT

In this paper, feature selection technique (FST), namely particle swarm optimization (PSO), has been used. The filter-based PSO is a search method with correlation-based feature selection (CBFS) as a fitness function. The FST has two key goals of improving classification efficiency and reducing feature counts. Artificial neural network (ANN)-based multilayer perceptron network (MLP) and deep learning (DL) have been considered the classification methods on two benchmark autism spectrum disorder (ASD) datasets. The experimental result was compared to the non-reduced features and reduced feature of ASD datasets. The reduced feature gives up enhanced results in both classifiers MLP and DL. In addition, an experimental study on the exhibitions of these methodologies has been conducted. Finally, a new trend of PSO-MLP and PSO-DL-based classification model is proposed.

KEYWORDS

Deep Learning (DL), Feature Selection Technique (FST), Multilayer Perceptron Network (MLP), Particle Swarm Optimization (PSO)

1. INTRODUCTION:

Autistic Spectrum Disorder (ASD) is not only detrimental to the afflicted individual but also to the patient’s family and but also society at large. Several researchers focusing in the area of behavioral science and computational intelligence have recently worked on ASD. They worked intelligent models like ANN and DNN to detect and diagnosis the ASD. One of the psychiatric problems is Autistic Spectrum Disorder (ASD). ASD is a psychological disorder that limits the use of linguistic, communicative, cognitive, capabilities as properly as social competencies and abilities. In recent times, ASD has been studied in the behavioural sciences using intelligent technique based around machine learning to speed up the screening time or to improve sensitivity, accuracy of the diagnosis process (Thabtah, 2017). The ASD is a group of Neuro-developmental disabilities disorders that cannot be cured but may be ameliorated by early detection with the help of Machine Learning (ML). To deal with the ASD problem, this research paper a ML structure used for ASD screening applying a new class of learning called Deep Learning (DL). This research objective to enhance the ASD screening procedure by including a new DL based on Artificial Neural Network (ANN). The ANN
based models have showed their advantage in various classification field such as text classification, image classification, and medical diagnosis. Feature selection was a technique used for data pre-processing and it was ideal to implement it earlier than fitting model in machine learning. Selection by mean here was by selecting features and attributes or feature subset in a dataset to be fit into model and tested for performance. In this we used the meta-heuristic search based method called Particle Swarm Optimization (PSO) it is feature Selection technique (FST). The proposed model MLP, DL and PSO-MLP and PSO-DL is helpful for identification and detection of ASD problems and also important features are recognized in this study.

This paper contributes in three ways

1. It examines the performance of the MLP and DNN classifiers in terms of accuracy, sensitivity, specificity, and F1-score.
2. The proposed PSO-CBFS has used for feature optimization.
3. It compares and analyzes the performance of classifiers MLP and DNN classifiers with optimized features by PSO-CBFS.

2. RELATED WORKS:

Many researchers have been working in the field of computer-based disease diagnosis systems. The results of the studies show the usefulness of data mining models in ASD-related problems.

Thabtah & Peebles (2020) worked on the classification algorithm like PRISM, CART, AdaBoost, Bagging, Nnge, RIDOR, Rules-Machine Learning (RML), RIPPER and C4.5 for classification of 3 Autism Spectrum Disorder children, adolescents, and adults. The result shows that the Rules-Machine Learning (RML) performed better than other models (Thabtah & Peebles, 2020). Misman et al., (2019) worked on the Support Vector Machine (SVM) and Deep Neural Network (DNN) for detection of two ASD adult datasets. The outcome shows that the DNN highest accuracy obtained by 99.40% on the first dataset and achieved 96.08% on the second dataset compared to SVM (Misman et al., 2019). Wang et al.(2019) worked on the feature engineering and feature encoding techniques, along with a deep learning classifier for ASD screening. The outcome shows that the achieved the 99% sensitivity and 99% specificity (H. Wang, Li, Chi, & Zhao, 2019). Hassan & Mokhtar (2019) worked on decision tree method were utilized to analyze related factors in datasets obtained from the National Database for Autism Research (NDAR) consisting of nearly 3000 individuals. The Decision Tree Classifier from scikit-learn, ADTree, CDT, J48, and LADTree classifiers have used, The experimental outcome they obtained the accuracy 90% using the decision tree classifiers (Hassan & Mokhtar, 2019). Diabat et al. (2019) worked on the classifiers C4.5, PART, RIPPER and Voted Perceptron and Ensemble Classification for Autism Screening (ECAS). The outcome shows that the Ensemble Classification obtained the highest accuracy 100% compared to other models (Diabat & Al-shanableh, 2019). Abdullah, Rijal, & Dash (2019) worked on feature selection technique Chi-square and Least Absolute Shrinkage and Selection Operator (LASSO) and classification algorithm Random Forest, Logistic Regression and K-Nearest Neighbors. The outcome shows that the Logistic Regression scored the highest accuracy 97.541% with 13 selected features based on Chi-square FST (Abdullah, Rijal, & Dash, 2019). Akyol, Gultepe, & Karaci (2018) proposed model like Fuzzy Rule (FR) and Combination of Logistic Regression wit Fuzzy Rule (LR-FR) for detection of ASD problems. The outcome shows that the LR-FR obtained the accuracy 97.33% was better than FR (Akyol, Gultepe, & Karaci, 2018). Vaishali & Sasikala (2017) proposed Binary Firefly Algorithm (BFA) feature selection wrapper and classifiers Naive Bayes, J48 Decision Tree, SVM, K-NN and MLP used for ASD detection. The experimental outcome shows that the MLP and SVM have obtained the highest accuracy 99.66% with selected feature by BFA algorithm compared to other models (Vaishali & Sasikala, 2017).
3. PROPOSED FRAMEWORK

The figure 1 shows the systematic a representation of “A proposed framework”.

In this proposed framework, have the following several steps.

3.1. Dataset Used

In this research work have used two kind of chronic disease dataset namely Autistic Spectrum Disorder Screening Data for Adolescent (ASDS Adolescent) and Autistic Spectrum Disorder Screening Data for Children (ASDS Child) datasets.

Autistic Spectrum Disorder Screening Data for Adolescent: The Autistic Spectrum Disorder Screening Data for Adolescent (ASDS Adolescent) data set is collected from UCI repository having autism screening of adults that contained 20 features with one class label and the total number Instances 104.

Autistic Spectrum Disorder Screening Data for Children Data Set: The Autistic Spectrum Disorder Screening Data for children (ASDS Child) data set is collected from UCI repository having autism screening of adults that contained 20 features with one class label and the total number Instances 292.

3.2. Feature Selection Technique

Particle swarms optimization (PSO): Particle Swarm Optimization (PSO) is a meta-heuristic because it gives little or no indication of problem is optimization and can search for very large locations of
candidate solutions (Blum & Roli, 2001). PSO is a population-based stochastic optimization technique developed by Dr. Eberhart and Dr. Kennedy in 1995 and inspired by bird flocking or fish schooling social behaviors (Kennedy & Eberhart, 1995; Shi & Eberhart, 1998). These swarms conform to a cooperative way to find food, and each member in the swarms continue shifting the search pattern according to the learning experiences of its own and other members (D. Wang, Tan, & Liu, 2018). We used the CBFS feature evaluator with the PSO search method to select features for ASD datasets. The CBFS evaluates the worth of a subset of attributes by considering the individual predictive ability of each feature along with the degree of redundancy between them. Subsets of features that are highly correlated with the class while having low inters correlation are preferred. The PSO search investigated the attribute space using the Particle Swarm Optimization (PSO) algorithm.

The Steps of proposed PSO algorithm

**Step 1:** Begin input ASD datasets

**Step 2:** Tuned the PSO parameter to find the best subset feature
go to step 3 and Step 4

**Step 3:** Choose the basic parameter of PSO
Set seed size
Set size of particles in the swarm
Set report frequency
Set iterations
Set best fitness
Set social weight
Set individual weight
Set mutation type
Set inertia weight
Set a start point for the search of particle
Set probability of mutation

**Step 4:** Choose the EV as CBFS
Set learning algorithm CBFS
Set size of thread pool
Set debugging info
Set missing separate
Set the size of the thread pool
Set check evaluator capabilities
Set pre-compute the full correlation matrix
Set identify locally predictive attributes

**Step 5:** Run the PSO-CBFS under the selected and set parameter

**Step 6:** Find the best feature subset of ASD datasets

**Step 7:** End

### 3.3. Classification Techniques

**Multilayer Perceptron Network (MLP):** A Classifier that uses back propagation to classify instances. This network can be built by hand, created by an algorithm or both. The network can also be monitored and modified during training time. The nodes in this network are all sigmoid (except for when the class is numeric in which case the output nodes become unthresholded linear units). In the below figure 2 shows the proposed MLP.

**Deep Learning (DL):** Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction. Deep learning discovers intricate structure in large data sets by using the back-propagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation
in each layer from the representation in the previous layer (Lecun, Bengio, & Hinton, 2015). Deep Neural Network (DNN) and Neural Network (NN) are intriguing issues of pattern identification and machine learning investigation in the present day have been connected to various tasks in a variety of areas. The best applications that bargain with it have originated from the medicinal services and healthcare areas. Deep Neural Network (DNN) procedures and frameworks that become familiar with their parameters as well as appropriate features: usually a hierarchy of features which increasingly period the semantic gap between low-level data and high-level label assignments. In the below figure 3 shows the proposed DNN diagram.

### 3.4. Evaluation Metrics

One significant step is the data mining analysis of the results calculation using evaluation metrics. Through using such metrics we can analyze and compare the performance of the proposed models. For this purpose we have used four metrics, which are accuracy, specificity, sensitivity, F1 -score calculated using the True Negative (TN), False Positive (FP), False Negative (FN) and True Positive (TP) rates.

As refer to above, four metrics have been used to calculate the performance of proposed model. We can be calculated using Equations. (1)– (4).

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}
\]

\[
\text{Sensitivity} = \frac{TP}{TP + FN} \tag{2}
\]

\[
\text{Specificity} = \frac{TN}{TN + FP} \tag{3}
\]

\[
\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}
\]
Specificity = \frac{TN}{TN + FP} \quad \text{(3)}

F1-Score = \frac{2TP}{2TP + FP + FN} \quad \text{(4)}

Finally, we compare the performance of proposed models MLP and DL to PSO-MLP and PSO-DL.

4. RESULT AND DISCUSSION

In this study, we used the concept of ANN method with PSO.

PSO Search explores the attribute space using the Particle Swarm Optimization (PSO) algorithm. The table 1 shows the feature selected by the proposed PSO in both ASD datasets.

Table 1 demonstrates how many features selected by our proposed FST PSO for two ASD datasets. The PSO has reduced the 11 features out of the 20 features of ASDS Adolescent dataset. Then selected the 9 important features of ASDS Adolescent dataset. The PSO has reduced the 09 features out of the 20 features of ASDS Child dataset. Then selected the 11 important features of ASDS Child dataset.

Table 1. Features selected by proposed PSO

| Particle Swarm Optimization (PSO) | Name of Dataset    | Total Selected Features | Features Detail |
|----------------------------------|--------------------|-------------------------|-----------------|
|                                  | ASDS Adolescent    | 09                      | \(F_1, F_2, F_3, F_4, F_5, F_6, F_7, F_8, F_{10}, F_9, F_{18}\) |
|                                  | ASDS Child         | 11                      | \(F_1, F_2, F_3, F_4, F_5, F_6, F_7, F_8, F_9, F_{10}, F_{18}\) |
Accuracy is the percentage of correctly classified instances among all instances. The accuracy of the proposed models used in these experiments is reveal in Table 2.

Table 2. Accuracy of classifications model (In %)

| Name of Algorithms | ASDS Adolescent | ASDS child |
|--------------------|-----------------|------------|
| MLP                | 89.4231         | 99.6575    |
| DL                 | 89.4231         | 95.2055    |
| PSO-MLP            | 91.3462         | 100.00     |
| PSO-DL             | 98.0769         | 100.00     |

Table 2 shows the accuracy of classifiers with reduced and all features of ASDS datasets. The all features of ASDS Adolescent dataset the MLP and DL obtained the equal accuracy (89.4231%). In case of reduced features of ASDS Adolescent the PSO-MLP obtained the accuracy (91.3462%) and PSO-DL obtained the accuracy (98.0769%). The comparison of PSO-MLP and MLP in ASDS Adolescent dataset the PSO-MLP obtained the higher accuracy (1.9231%). The same way comparison of PSO-DL and DL in ASDS Adolescent dataset the PSO-DL obtained the higher accuracy (8.6538%). Similarly the all features of ASDS Child dataset the MLP obtained the accuracy (99.6575%) and DL obtained the accuracy (95.2055%). It means the ASDS Child dataset with all features the MLP obtained the higher accuracy (4.452%) compared to the DL. In case of reduced features of ASDS Child dataset the PSO-MLP obtained the accuracy (100%) and PSO-DL obtained the accuracy (100%). The comparison of PSO-MLP and MLP in ASDS Child dataset the PSO-MLP obtained the
higher accuracy (0.3425%). The same way comparison of PSO- DL and DL in ASDS Child dataset the PSO- DL obtained the higher accuracy (4.7945%).

The figure 4 shows the accuracy graph of our proposed model in case ASDS Adolescent dataset. The graph is clearly shows that the all feature and reduced feature of ASDS Adolescent dataset. The reduced dataset with proposed classification models is better performed as compared to the all feature with proposed classification models. In all the cases the most suitable model is PSO-DL because it obtained the highest point in the graphs.

The figure 5 shows the accuracy graph of our proposed model in case ASDS Child dataset. The graph is clearly shows that the all feature and reduced feature of ASDS Child dataset. The reduced dataset with proposed classification models is better performed as compared to the all feature with proposed classification models. In all the cases the most suitable model is PSO-MLP and PSO-DL because it obtained the highest point in the graphs.

Sensitivity analysis techniques measure the rate of change at the output of a sample due to adjustments in the input of a variable. The sensitivity of the proposed models used in these experiments is reveal in Table 3.

| Name of Algorithms | ASDS Adolescent | ASDS Child |
|--------------------|-----------------|------------|
| MLP                | 85.36           | 99.33      |
| DL                 | 92.68           | 96.02      |
| PSO-MLP            | 85.36           | 100        |
| PSO-DL             | 97.56           | 100        |

Table 3. Sensitivity of classifications model (In %)
The data shows the sensitivity of proposed model in both ASD datasets. The proposed MLP and PSO-MLP has obtained equal sensitivity 85.36% of ASDS Adolescent dataset. While the proposed DL obtained the sensitivity 92.68% and PSO- DL obtained the sensitivity 97.56% of ASDS Adolescent dataset. The PSO-DL obtained the higher sensitivity in all models in ASDS Adolescent dataset. The same way proposed MLP obtained the sensitivity 99.33% and PSO-MLP obtained sensitivity 100% in ASDS child dataset. While the proposed DL obtained the sensitivity 96.02% and PSO- DL obtained the sensitivity 100% of ASDS child dataset. The PSO-MLP and PSO-DL obtained the higher sensitivity in all models in ASDS child dataset.

Specificity analysis is the proportion of actual negatives that are correctly identified. The specificity of the proposed models used in these experiments is reveal in Table 4.

Table 4. Specificity of classifications model (In %)

| Name of Algorithms | ASDS Adolescent | ASDS child |
|--------------------|-----------------|------------|
| MLP                | 92.06           | 100        |
| DL                 | 87.30           | 94.32      |
| PSO-MLP            | 95.23           | 100        |
| PSO-DL             | 98.41           | 100        |

Table 4 shows the specificity of the MLP, DL, PSO-MLP and PSO-DL classifiers in both ASDS datasets as compared to each other. In the case of ASDS Adolescent dataset the PSO-DL has the best results. Whereas in the case of ASDS child dataset the MLP, PSO-MLP and PSO-DL have provide the highest results compared with the DL.

F-measure (F1-score) has been introduced to balance between sensitivity and specificity. The F1-score of the proposed models used in these experiments is reveal in Table 5.

Table 5. F1-score of classifications model (In %)

| Name of Algorithms | ASDS Adolescent | ASDS child |
|--------------------|-----------------|------------|
| MLP                | 86.41           | 99.66      |
| DL                 | 87.35           | 95.39      |
| PSO-MLP            | 88.60           | 100        |
| PSO-DL             | 97.56           | 100        |

5. CONCLUSION:

Autistic Spectrum Disorder (ASD) is a psychological disorder that limits the use of linguistic, communicative, cognitive, capabilities as properly as social competencies and abilities. Early identification of this disorder can significantly decrease the health issues thus it enhance the overall
mental health of suffering peoples. We used the Machine learning technique for identification of ASD problems. The FST like PSO has used for identification of the important features of ASD. The classification techniques namely MLP and DL have used for detection of ASD problems. The proposed models have been experimented implemented with 2 ASD datasets. The result the classification MLP and DL gives 89.42% and 89.42% accuracy with ASDS Adolescent dataset. Similarly the classification PSO-MLP and PSO-DL gives 91.35% and 98.08% accuracy with ASDS Adolescent dataset. The classification MLP and DL gives 99.66% and 95.20% accuracy with ASDS child dataset. Similarly the classification PSO-MLP and PSO-DL gives 100% accuracy with ASDS child dataset. The experimental outcome clearly shows that the PSO-MLP and PSO-DL enhance the classification performance of the applied classifiers for ASD datasets.

The number of features and classification techniques will play an important role in further classification of CAD.
REFERENCE

Abdullah, A. A., Rijal, S., & Dash, S. R. (2019). Evaluation on Machine Learning Algorithms for Classification of Autism Spectrum Disorder (ASD). In Journal of Physics: Conference Series (pp. 1–6). IOP Publishing. doi:10.1088/1742-6596/1372/1/012052

Akyol, K., Gultepe, Y., & Karaci, A. (2018). A Study on Autistic Spectrum Disorder for Children Based on Feature Selection and Fuzzy. International Congress on Engineering and Life Science, 804–807.

Al Diabat, M., & Al-shanableh, N. (2019). Ensemble learning model for screening. International Journal of Computer Science and Information Technologies, 11(2), 13–14. doi:10.5121/ijcsit.2019.11205

Blum, C., & Roli, A. (2001). Metaheuristics in Combinatorial Optimization: Overview and Conceptual Comparison ACM. Computer Survey, 35(3), 268–308. Advance online publication. doi:10.1145/937503.937505

Hassan, M. M., & Mokhtar, H. M. O. (2019). Informatics in Medicine Unlocked Investigating autism etiology and heterogeneity by decision tree algorithm. Informatics in Medicine Unlocked, 16(July), 100215. doi:10.1016/j.imu.2019.100215

Kennedy, J., & Eberhart, R. (1995). Particle Swarm Optimization. Academic Press.

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. 10.1038/14539

Misman, M. F., Samah, A. A., Ezudin, F. A., Majid, H. A., Shah, Z. A., Hashim, H., & Materials, A. (2019). Classification of Adults with Autism Spectrum Disorder using Deep Neural Network. University of Edinburgh.

Shi, Y., & Eberhart, R. (1998). A Modified Particle Swarm Optimizer. Academic Press.

Thabtah, F. (2017). Autism Spectrum Disorder Screening: Machine Learning Adaptation and DSM-5 Fulfillment. Academic Press.

Thabtah, F., & Peebles, D. (2020). A new machine learning model based on induction of rules for autism detection. Health Informatics Journal, 26(1), 265–286. doi:10.1177/1460458218824711 PMID:30693818

Vaishali, R., & Sasikala, R. (2017). A machine learning based approach to classify Autism with optimum behaviour sets. IACSIT International Journal of Engineering and Technology, 5(x), 1–6.

Wang, D., Tan, D., & Liu, L. (2018). Particle swarm optimization algorithm: An overview. 10.1007/s00500-016-2474-6

Wang, H., Li, L., Chi, L., & Zhao, Z. (2019). Autism Screening Using Deep Embedding Representation. In Lecture Notes in Computer Science. Springer.

Sanat Kumar Sahu is working as Assistant Professor in Department of Computer Science, Govt. Kaktiya PG College, Jagdalpur (Bastar) Chhattisgarh. He has received his Master’s Degree in Computer Application from Guru Ghasidas Vishwavidyalaya, Bilaspur in 2008 and M. Phil in Computer Science from Dr. C.V. Raman University, Kota, Bilaspur in 2009. He has more than 7 years teaching and 02 years research experience. He has published more than 12 research paper in reputed journals and attended workshop and conference at national and international level. His area of interest includes soft computing, Machine learning, and data mining.

Pratibha Verma is currently a Ph.D. candidate in Department of Information Technology at Dr. C.V. Raman University, Bilaspur (Chhattisgarh), India. She received her master’s Degree in Computer Application from Chhattisgarh Swami Vivekananda Technical University Bhilai (Chhattisgarh), India in 2012. Her research interests include machine learning and data mining.