An Improved Feature Selection (IFS) Algorithm for Detecting Autistic Children Learning Skills

M.S. Mythili and A.R. Mohamed Shanavas

¹Department of Computer Applications, Bishop Heber College,
Tiruchirappalli - 620 017, TamilNadu, India.
²Department of Computer Science, Jamal Mohamed College,
Tiruchirappalli - 620 020, TamilNadu, India.

Children with autism spectrum disorder follow a special process pattern from other children and develop at a special rate. Educating children with autism is expounded to intensive endeavor, involving a team of professionals of assorted special instruction and therapies to subsume children behavioral and activity, biological process, social and academic needs. Students with autism usually need explicit teaching across a range of settings to generalize skills. The knowledge and information gained from analysis helps the parents and school teachers to convey decent learning surroundings for the autism children. Feature Selection is a very important topic in data mining, particularly for high dimensional datasets. Feature Selection is a method usually employed in machine learning, whereby subsets of the options accessible from the data are described for the application of a learning algorithm. The Main purpose of this paper is to propose the Improved Feature Selection algorithm. The Improved algorithm is developed by combining the filters and wrappers. The Correlation Based Feature Selection (CFS) with the best first search act as a filter for removing impertinent options. Wrapper Subset Evaluator with the best first search is employed as a wrapper and it absolutely reduces the redundant options. It is accustomed to improve the accuracy of the classification for the autism children by analyzing the four totally different classifiers such as SVM, *J48, Multilayer* Perceptron and IB1are *used.* The overall purpose of this paper work is to propose the foremost effective algorithmic program for autism children with sensible potency and accuracy.

Key words: Data Mining, Classification, Filter Approach, Feature Selection Algorithms.

The data mining is customary recently as a result of all the information is captured with every dealing by the transaction process method and thus the system's capability is extremely store huge data and therefore the measurement is among the vary of terabytes. Data Mining is that the method of discovering new vital correlations, non-trivial, implicit, antecedently unknown and potentially helpful information from large volumes of knowledgeable data. Society produces large amounts of raw data from varied sectors together with business, science, medicine, web mining. Data mining uses applied mathematics or a mathematical technique to bring out the relationships hidden among the informational data. Feature Selection is that the method of choosing the set of relevant features by removing redundant, irrelevant and clanging information data from the initial dataset. Feature Selection methods comprise two categories: Filter approach and Wrapper approach.

* To whom all correspondence should be addressed.
E-mail: mythhariniee24@yahoo.com

In filter approach, the features were supported based on criteria which are independent of the actual learning algorithmic rule to be applied to the data. In wrapper approach, the Feature Selection is predicated on a wrapper, which may be a set of attributes and are evaluated with a learning algorithmic rule. Feature Selection can be divided in filter methods and wrapper methods. Filter methods is outlined as using some actual property of the data so as to pick the feature using the classification algorithmic rule. Entropy measure has been used as filter technique for feature selection for classification. Wrapper approaches apply classification algorithmic rule to every candidate feature subset then measure the feature subset by threshold functions that utilize the classification result. This paper is structured as follows. Section two discusses the Autism children role with their signs and symptoms. Section three describes the related work. Section four elaborates and presents the idea of Classification and discusses the aspects of classification algorithmic rule. Section five enumerates a Data Preprocessing. Section six explains the Implementation of model construction. Section seven narrates the idea of filter approach with the wrappers. Section eight clearly explains the proposed methodology with the algorithm results and discussions. Section nine accurately elaborates the results and discussion via in table and in the graphical representation. Section ten gives the conclusion and future work.

AUTISM

Autism Spectrum Disorder (ASD) is a Pervasive Developmental Disorder (PDD). It is a bio-neurological developmental disability typically showing before the age of three, best renowned for impairing a child's ability to speak and move. Life-long disabilities considerably impact in many areas of development: communication impairments, social issue, sensory process deficits and a necessity for solid routines among their lives. Characteristics of Autism manifests in an exceedingly myriad of ways: delay in verbal development, a necessity to end what they start, a rather h3 resistance to vary in daily routine, lack of spontaneousness, distress at being touched and also the ability to indicate any reasonably feeling of emotion, as well as an inability to process their method and reply to humor.

Literature review

Jen-Da shie, Shyi-Ming Chen, et al has bestowed a technique for dealing with Feature Subset Selection supported on fuzzy entropy measure for handling classification issues. It considers boundary samples whereas choosing options. Thus fuzzy entropy technique, measures the impurity of a feature. This Feature Selection technique selects relevant features to urge higher average classification accuracy. It does not measure the purity of datasets. Behrouz Minaei, Hossein Ghaffarian, Hamid Parvin, et al has elaborated a process for Innovative feature subset selection with the fuzzy entropy measure for handling classification issues. The entire dataset is classified and according to silhouette value, the most effective variety of clusters in the dataset is found. S.Sethuramalingam, Dr.E.R.Naganathan, et al has enumerated a method on Hybrid Feature selection for network Intrusion. A new algorithm proposed on hybrid method to identify the significance of features. The hybrid method combines Information Gain and Genetic Algorithm to select features. Clustering is applied on elite options for classification. Rough set and neural network based reduction has been modified by Thangavel .K, & Pethalakshmi et al, describes the reduction attribute with the assistance of medical datasets. Protocol based classifications has been explained by Kun-Ming Yu, Ming-Feng Wu, and Wai-Tak Wong et al, which describes the protocol based classification by using genetic algorithmic with the logistic Regression as the method and enforced implemented by KDD 99 dataset. Data Analysis methodologies were delineated by,Shaik Akbar, Dr.K.Nageswara Rao, Dr.J.A.Chandula et al, deals with eleven data computing technique associated with IDS are divided groups into categories. Those strategies are based on computational Fuzzy logic and Bayesian networks, Artificial Intelligence, Expert Systems, agents and neural networks and other are biological ideas Genetics and Immune systems. Discernibility matrix was described by Chuzhou et al, gives a neat rationalization regarding the discernibility matrix function and reduction of features. Misuse and Anomaly detection exploitation using SVM, ANN approaches mentioned by T.Subbulakshmi et al, notifies the detection rate and warning rates. Multilayer

Perceptrons, NaïveBayes classifiers and Support vector machines with three kernel functions were used for sleuthing intruders. The Precision, Recall and F- Measure for all the technique were calculated.

Classification

Classification technique is a supervised learning technique and it plays an important role in data mining. It generates the models and very useful to classify the dataset. The technique is a two step process .The first process helps for the model construction and it builds a model from the training data. The second one is the model usage to check the accuracy of the model. It fully predicts the categorized class labels for all the samples data.

Classification Algorithms

The classifier algorithm deals with SVM, *J48, Multilayer* Perceptron and IB1algorithm .The classifiers are mentioned below

J48

The J48 classifier inherited all the concepts from ID3.The whole concepts of J48 classifier was developed by Quinlan Ross. It gets all the additional information from the Hunt's classifier. J48 splits the attribute into two levels and it is determined by the threshold values. The surrogate splitting is needed to check the missing values. It starts the classification of records from the test data by traversing the nodes from the root node to leaf node and it performs the conditional test .It also checks the attributes, the best attribute should be selected and made the attribute as the root node of the attribute. Attribute selection such as Gini index, Information gain helps in ranking the classification of the attributes. An attribute selection measure is used and analyses to construct the decision tree. The Information gain reduces the entropy value to the level which reduces the level of impure splitting. The main aim of pruning is to construct the optimal classifiers inside the decision tree to bolster the accuracy of classification.

Support Vector Machine

The Support Vector Machine is a supervised learning technique and it is easily associated with learning algorithms. It analyzes the data and acknowledges patterns. It is mainly used for classification and statistical analysis purpose. It is a non-probabilistic binary linear classifier to build a model and it assigns the new samples into one category to the other. SVM expeditiously perform a non-linear classification .SVM Classification is undergone by the kernel, which map their inputs into high-dimensional feature area. It plays an important role in hand written character recognition, face detection and it is widely applied in classification and regression.

Multilayer Perceptron

An artificial neural network model has the Input and output data and they are connected with the corresponding weights. In a directed graph, every layer is correctly connected to the other one. Multilayer Perceptron is the supervised learning technique and it is called as back propagation for training the network. The Multilayer Perceptron has three layers - input layer, output layer and the hidden *layers* of nodes. The main advantages are straightforward, easy to use, and they easily map input/output nodes.

IB1

It is a supervised learning classifier. It has the class labels for each record and with the help of the nearest neighbor classifier the distance is calculated. If there is not any class label for the records the mostly used label is used. It is a faster algorithm classifier and it sometimes is called as lazy classifier.

Attribute Identification

Dataset collected from autism children database consists of

Attributes	Description
Attention	Attention Skills of autism child
Hand writing	Hand Writing Skills of autism child
Spell	Spelling Skills of autism child
Lang	Language Skills of autism child
Reading	Reading Skills of autism child
Writing	Writing Skills of autism child
Interaction	Interaction Skills of autism child
Memory of child	Memory Skills of autism child
Listen	Listening Skills of autism child
Speaking	Speaking Skill of autism child s
Maths	Mathematical Skills of autism child
Science	Science Skills of autism child
Art	Art Skills of autism child
Music	Music Skills of autism child
physical	Physical Skills of autism child
Child class label	Levels of autism child(High,Medium, Mild)

Data preprocessing

Data preprocessing is an important one to eliminate the noisy data, missing data and redundant attributes over the Datasets. The Datasets utilized in the classification algorithmic rule applied the data preprocessing to remove the unwanted data's. The data have to be clear, correct and it is preprocessed for handling missing or redundant attributes. The data should be a complete and it should be consistent data to induce the best outcome from the Data Mining methodology.

Implementation of model construction

The Weka is the open provide software system tools and it is the integration of machine learning algorithms and the data preprocessing. It is widely utilized in data mining applications which contain the association rule mining, classification and clustering. It accepts only the ARFF (Attribute Relational File Format), CSV (Comma Separated Value) file. The detection of autism children learning skills has different attributes such as attention skills, Handwriting skills, spelling skills, language skills, Reading skills, writing skills, memory skills, listening skills.maths skills, science skills, music skills, physical skills, speaking skills, art skills, and the interaction skills, 200 instances were taken for the detection. The filter and the wrapper approaches have been implemented in Weka tool.

Generalized Filter and Wrapper Approaches

WEKA provides a supervised attribute filter that is versatile be flexible and permits varied search and analysis strategies to be combined. It makes use of two properties that are evaluator and search. Evaluator determines the factors used for choosing the attribute.

Different evaluators that can be used are:

Correlation-based feature selection

It considers the prophetic ability of every feature and finds out the relevancy. It conjointly checks for redundancy between the chosen options. Finally the set of attributes that are extremely correlated with the category and are less redundant is chosen.

$$Merit_{Sk} = \frac{K \overline{r_{cf}}}{\sqrt{k + k(k-1)\overline{r_{ff}}}}$$

Best First

It searches for the attribute set by greedy hill climbing method together with backtracking.

The backtracking relies on the conception that if some variety of consecutive nodes is found such they are not improve the performance then backtracking is completed. It should apply forward approach wherever it starts from empty set of attributes and goes on adding the consecutive one. It should conjointly choose backward approach wherever it starts from a collection of all attributes and removes one by one. It should conjointly adopt a midway between each approaches wherever search is completed in each directions that is additionally referred to as hybrid approach

Wrapper Subset Eval

Evaluates attribute sets by employing a learning scheme. Cross validation is employed to estimate the accuracy of the learning scheme for a collection of attributes.

Ranker

Individual evaluations of the attributes are done and that they are ranked consequently.

Proposed Algorithm

The proposed Algorithm is as follows:

Input: D (F_1, F_2.....F_n)// Training data with n features

Output: Best Subset with maximum Accuracy and minimum error rate

//**Feature selection**

1. Let S= set of n features F_1, F_2 ....F_n

2. Evaluate the subsets of features into S, where S is the feature Subset.

3. for each Feature Selection S,

4. Begin

5. Retrieved data=evaluate AS(S) // AS=Attribute Selection

// Compute the Filter process by

6. Perform the filter process to compute Correlation-based heuristic evaluation on Best feature subset

$$Merit_{Sk} = \frac{K \overline{r_{cf}}}{\sqrt{k + k(k-1)\overline{r_{ff}}}}$$

7. If Retrieved data>Initial data then

8. Initial data=Retrieved data //Keeping track of the Initial Subset encountered in the filter process.

9. Compute Accuracy

10. End

// Compute the Wrapper Process by

11. Choose a machine learning algorithm to evaluate the feature subset with best first search

and threshold value

12.   Final data=Initial data
13.   Repeat
14.   Begin
15.   within each Initial data
16.   End
17.   Until (Initial data>Final data)// Keeping track of the Final Subset in wrapper process
18.   Best Subset =Initial data
19. Compute Accuracy
//Classification Process
20. Run a 10-fold cross validation on the Best Subset by the classifiers such as SVM, *J48, Multilayer* Perceptron and IB1
21. Performance Evaluations gives the maximum Accuracy and minimum error rate
22. Return Best Subset

Proposed Methodology
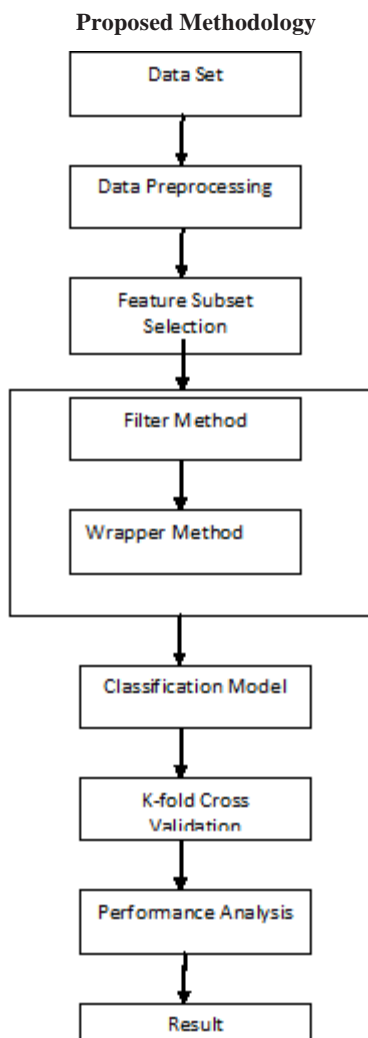
RESULTS AND DISCUSSION

The analysis and interpretation of classification is time consuming process that needs a deep understanding of statistics. The process needs a large amount of time to finish and expert analysis to look at the classification and relationships within the data.

In Table I the performance result of classifiers without the feature selection is studied. Hence Support Vector Machine with Normalized Polykernel can be termed as more accurate than other classifiers.

In Table II the performance result of classifiers with the feature selection is discussed. The proposed algorithm is applied to the dataset and all the classifiers performance (accuracy), is increased and the time, error rate is decreased. With the proposed algorithm Support Vector Machine with Normalized Polykernel can be termed as more accurate than other classifiers with 98%

The Figure 1 shows that the graphical representation of accuracy of the classifier algorithm. It clearly enumerates that with the feature selection all the classifiers algorithm accuracy has been increased. The proposed algorithm suggests that the Support Vector Machine is a very best classifier for analyzing the autism children
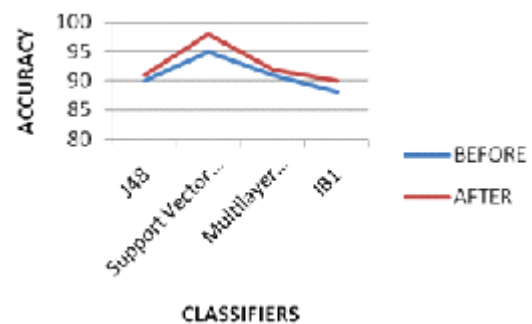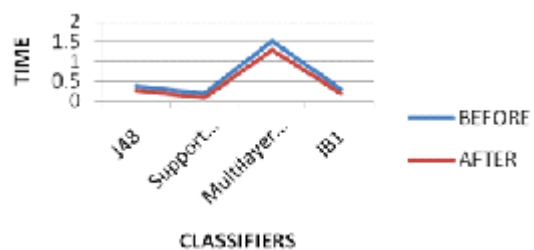

Fig. 1. Accuarcy of the classifier algorithm

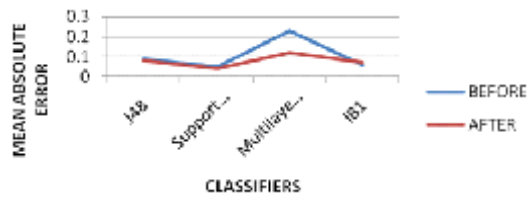**Fig .2.** Performance related to time

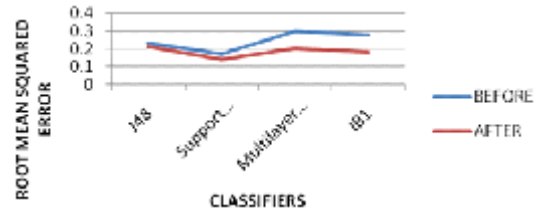**Fig. 3.** Mean absolute error rate of different classifiers

Fig. 4. Root mean squared error rate of different classifiers

Table 1. Performance result of classifiers before feature selection

Classifier	Correctly Classified Instances	Incorrectly Classified Instances	Mean Absolute Error	RootMean Squared Error	Time Taken(s)
J48	90	10	0.0937	0.2327	0.4
Support Vector Machine	95	5	0.0531	0.1712	0.2
Multilayer Perceptron	91	9	0.2389	0.3012	1.5
IB1	88	12	0.0833	0.2887	0.3

Table 2. Performance result of classifiers after feature selection

Classifier	Correctly Classified Instances	Incorrectly Classified Instances	Mean Absolute Error	RootMean Squared Error	Time Taken(s)
J48	91	9	0.0871	0.214	0.3
Support Vector Machine	98	2	0.0474	0.1453	0.1
Multilayer Perceptron	92	8	0.1244	0.2023	1.3
IB1	90	10	0.0781	0.1887	0.2

performance consuming less time coupled with good accuracy with 98%.

The Figure 2 elaborates the performance related to time by the various classification algorithms. It clearly deals that with the feature selection all the classifiers algorithm time has been decreased. The proposed algorithm performance with the Support Vector Machine is very fast for analyzing the results within 0.1 seconds

The Figure 3 compares Mean absolute error rate among completely different classifiers .All the classifiers got the lower error rate. Support Vector Machine has lower error rate compared to different classifiers after the feature selection.

The Figure 4 elaborates the representation of root mean squared error rate among completely different classifiers in a graphical way. All the classifiers got the lower error rate after the feature selection. The algorithm which

proposed concludes that Support Vector Machine has lower error rate compared to different classifiers.

CONCLUSION

This study is approached to find the simplest classification for the application of machine learning to autism detection with the Improved Feature Selection (IFS) algorithm. Completely different feature selection algorithm is employed with the combination of filters and wrappers to classify autism children in an efficient manner. After doing experimental work, it is clear that Support Vector Machine has highest classification accuracy, lowest error rate with less time as compared to different classifiers algorithm network. To boost the results the Feature Selection techniques is applied. The classifier algorithms are

applied to autism dataset by reducing its attributes and enforced using WEKA machine learning tool. The machine learning is a good methodology which can be utilized in the field of autism detection. In future, the proposal of new algorithm is to create a learning framework for the autism children in fuzzy cognitive map and in the Multilayer Perceptron Network with fuzzy inference rules to boost the performance.

REFERENCES

1. Harun Uguz, "A hybrid approach for text categorization by using x2 stastics, Principal Component Analysis and partical swam optimization", *Academic Journals,* (8) 37, PP 1818-1828, 2013.
2. Hyunjin Yoon et al. "Feature Subset Selection and feature ranking for multivariate time series", IEEE Transactions on Knowledge and Data engineering, Vol.17, no. 9, September 2005.
3. Kemal Polat, Salih Gunes, "Prediction of hepatitis disease based on principal component analysis and artificial immune recognition system", Applied Mathematics and Computation 189(2007) 1282-1291.
4. Tahseen A. Jilani et.al, "PCA-ANN classification of hepatitis- c patients", International Journal of Computer Applications, volume-14-No.7, Febraury 2011.
5. Hung lian, "on feature selection with principal component analysis for one-class SVM", Pattern Recognition Letters 33(2012) 1027-1031.
6. Hlip-Ling Chen Et.al, "A new hybrid method based on local fisher discriminant analysis and support vector machines for hepatitis disease diagnosis" . Expert Systems with Applications 38(2011) 11796-11803.
7. Yilmaz Kaya, Murat Uyar, " A hybrid decision support system based on rough set and extreme learning machine for diagnosis of hepatitis disease". Applied Soft Computing 13(2013) .
8. Huda yasin et al., "hepatitis-c classification using data mining techniques", International Journal of Computer Applications, Volume 24, No.3, June 2011.
9. Javad salami sartakhiti, mohammad Hossein Zangooei, Kourosh Mozafari, " Hepatitis disease diagnosis using a novel hybrid method based on support vector machine and simulated annealing(SVM-SA)'', Computer methods and programs in Biomedicine 108 (2012)570-579.
10. Diana Dumitru , "Prediction of recurrent events in breast cancer using the Naive Bayesian classification", *Annals of University ofCraiova* Math. Comp. Sci. Ser. Volume 36(2), 2009.
11. M.S. Mythili, A.R.Mohamed Shanavas,"A Study on Autism Spectrum Disorders using classification Techniques",IJSCE,ISSN:2231-2307,Volume 4,Issue 5,Nov 2014.
12. M.S. Mythili, A.R.Mohamed Shanavas,"A Novel Approach to predict the learning skills of Autistic children using SVM and Decision Tree, IJCSIT, ISSN: 0975-9646, vol5 (6), 2014, 7288-7291