IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-6, Special Issue-2, March 2018 E-ISSN: 2347-2693

Comparative Study on the Feature Selection Techniques for Autism Spectrum Disorder

S. Padmapriya[1*], S. Murugan[2]

[1*]Department of Computer Science,Srimad Andavan Arts and Science College, Trichy, Tamilnadu, India
[2] Department of Computer Science, Nehru Memorial College, Puthanampatti, Tamilnadu, India

Available online at: www.ijcseonline.org

***Abstract*-**There is a burgeoning need to consider new ways of providing early education services for young and often newly diagnosed children with Autism Spectrum Disorder (ASD) and their families. Such children do not respond naturally to direct curricular delivery, typically utilized in inclusive classrooms that predominate public education, but instead, need an educational model incorporating intra and interpersonal development skills. Also, there is an essential need for the facility to keep track of and addressing uneven progress in specific areas; characteristic of learners with ASD. In this paper, ranking feature selection techniques like Information Gain, Chi-Square, Gain Ratio, ReliefF are used for pre-processing the ASD dataset.

I. INTRODUCTION

Autism spectrum disorders (ASD) [1] are a group of related brain-based disorders that affect a child's behavior, social, and communication skills. They include 3 of 5 disorders known as pervasive developmental disorders (PDDs) [3]. These are Asperger Syndrome, PDD-not otherwise specified and autistic disorder. Because maximum children with ASD will learn the early skills such as crawling, walking on time and sitting, lags in communication and social skills may not be as apparent to parents. Looking back, several parents of children with ASD can think of specific examples that suggest something was different, but nothing was indicating a dangerous problem.

Autism spectrum disorders are lifelong conditions with no known cure [2]. However, children with ASD can progress developmentally and learn new skills. Some children may develop so much that they no longer meet the standards for ASD, although milder symptoms may often persist.

About the delivery of education to students with special learning needs, the most recent global question was whether or not a full inclusion model was efficacious. However, since it is already happening that matter is controversial, and the more salient issue now has become how educators, parents, and other professionals can work collaboratively to ensure that it has done well. The model of full inclusion is one in which all students are to be considered as entirely functioning and fully valued members of the school community, with the primary responsibility for education has been placed on the regular classroom teacher. In most instances, this is a defensible ideology and a workable model. However, in regards to students with Autism Spectrum Disorder, the challenges require an extraordinary level of intervention. The purpose of this paper has to recommend the manner in which the match of special education and technology might lead to improved services for students, less stress for teachers, and a more positive learning outcome in general.

II. IMPORTANCE OF FEATURE SELECTION

The number of high-dimensional data that endures and is publically accessible on the internet has very developed in the past few years. Therefore, machine learning techniques have the challenge in dealing with the significant number of input features, which is modeling an attractive issue for researchers. To utilize machine learning techniques efficiently, preprocessing of the data is essential. Feature selection [5] is one of the most frequent and prominent methods in data preprocessing, and has become a necessary component of the machine learning process is also known as variable selection, attribute selection, or variable subset selection in machine learning and statistics. It is the method of removing irrelevant and detecting relevant features, noisy data or redundant. This technique speeds up data mining algorithms, enhances comprehensibility and predictive accuracy. Unrelated features are those that give no useful information, and irrelevant features provide no more information than the currently selected features. Regarding supervised inductive learning, feature selection presents a set of candidate features using one of the three approaches [6].

- The exact size of the subset of features that optimizes an evaluation measure.
- The smaller size of the subset that satisfies a certain restriction on evaluation measures.
- In general, the subset with the best commitment among size and evaluation measure.

In the process of feature selection [7], noise or redundant features in the data may be hinder in many circumstances, because they are not essential and relevant for the class concept such as microarray data analysis. When the number of samples is much less than the features, then machine learning gets particularly difficult, because the search space will be sparsely populated. Therefore, the model will not able to differentiate accurately between noise and relevant data. There are two major approaches to feature selection. The first is Individual Evaluation, and the second is Subset Evaluation. Ranking of the features is known as Individual Evaluation. In Individual Evaluation, the weight of an individual feature is assigned according to its degree of relevance. In Subset Evaluation, candidate feature subsets are constructed using search strategy.

III. FEATURE SELECTION TECHNIQUES

A. Information Gain Feature Selection Method

Entropy is commonly used in the information theory measure, which characterizes the purity of an arbitrary collection of examples. It is in the foundation of Gain Ratio, Information Gain [8] and Similarity Uncertainity (SU). The entropy measure is considered a measure of the system's unpredictability. The entropy of Y is

$$H(Y) = \sum_{y \in Y} p(y) \log_2(p(y)) \quad (1)$$

where *p(y)* is the marginal probability density function for the random variable *Y*. If the observed values of *Y* in the training data set *S* are partitioned according to the values of a second feature *X*, and the entropy of *Y* with respect to the partitions induced by *X* is less than the entropy of *Y* prior to partitioning, then there is a relationship between features *Y* and *X*. The entropy of *Y* after observing *X* is then:

$$H(Y|X) = \sum_{x \in X} p(x) \sum_{y \in Y} p(y|x) \log_2(p(y|x)) \quad (2)$$

where *p(y |x)* is the conditional probability of *y* given *x*.

Given the entropy is a criterion of impurity in a training set *S*, we can define a measure reflecting additional information about *Y* provided by *X* that represents the amount by which the entropy of *Y* decreases. This measure is known as IG. It is given by

$$IG = H(Y) - H(Y|X) = H(X) - H(X|Y) \quad (3)$$

IG [9] is a symmetrical measure, and it is given by equation (3). The information gained about *Y* after observing *X* is equal to the information gained about *X* after observing *Y*. A weakness of the IG criterion is that it is biased in favor of features with more values even when they are not more informative.

B. Chi-Square Feature Selection Method

Feature Selection via chi-square χ^2 test [8] is another, very commonly used the method. Chi-squared attribute evaluation evaluates the worth of a feature by computing the value of the chi-squared statistic concerning the class. The initial hypothesis H_0 is the assumption that the two features are unrelated, and it is tested by the chi-squared formula:

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \left(\frac{O_{ij} - E_{ij}}{E_{ij}} \right)^2 \quad (4)$$

Where O_{ij} is the observed frequency, and E_{ij} is the expected (theoretical) frequency, asserted by the null hypothesis. The greater the value of χ^2, the greater the evidence against the hypothesis H_0.

C. Gain Ratio Feature Selection Method

The Gain Ratio [8] is the non-symmetrical measure that is introduced to compensate for the bias of the Information Gain (IG) [7]. GR is given by

$$GR = \frac{Information\ Gain\ (IG)}{H(X)} \quad (3)$$

Information Gain (IG) is an equal measure.

$$IG = H(Y) - H(Y|X) = H(X) - H(X|Y) \quad (4)$$

The information gained about Y after scrutinizing X is similar to the information gained about X after examining Y. A delicacy of the IG measure is that it is predetermined in support of features with higher values even when they are not more informative.

As in the equation (3) presents, when the variable Y has prognosticated, then normalize the IG by splitting by the entropy of *X*, and vice versa. Because of this normalization, the GR values constantly come in the range [0, 1]. A value of GR = 1 indicates that the knowledge of *X* completely predicts *Y*, and GR = 0 means that there is no relation between *Y* and *X*. In opposition to IG, the GR favors variables with fewer values.

D. ReliefF Feature Selection Method

ReliefF [10] is a sort of ranking algorithm that attempts to assign a fair rank to each feature iteratively. The algorithm in the first step considers a zero vector based on the number of features. Then in each step, the algorithm selects two samples, one of them must be the nearest neighbor with respect to the class of the selected sample and the other must be the nearest sample from the other class, they update the aforementioned vector in each iteration. The algorithm will run m times, where m is lower than the sample size.

IV. DESCRIPTION ABOUT THE DATASET

The autistic child dataset was collected from various special schools in and around Trichy. The following table 1

represents the description of the features that have been collected from schools.

TABLE 1: FEATURE NAMES AND ITS DESCRIPTION IN THE ASD DATASET

Sl.NO	Feature Name	Description of the Feature
1	Before speech loss (SL)	Yes/No (1-0)
2	Obstetric risk factors ≥ 3 (in 90 ADSL)	Of or relating to or used in or practicing obstetrics risk factors
3	Psychosocial event before SL (S1=1,S2=2,S3=3,S4=4)	S indicates percentage of effectiveness
4	During SL (for 6 months or less after onset)(D1=2,D2=4,D3=6))	D1=1-2,D2=2-4,D3=4-6(months)
5	verbal/non-verbal communication	-
6	stereotyped behaviors	Like normal person (Y/N) – 1/0
7	Lack of empathy(L-1:M-2:H-3)	Understanding and entering into another's feelings (l-low; M-Medium; H-High).
8	adaptive behaviors	Having a capacity for adaptation
9	Head-growth decelerations	Head-growth decrease in rate of change
10	interact with other people	Yes/No (1-0)
11	Fearfulness	Yes/No (1-0)
12	At first visit (Fear-1;Less Fear-0)	Yes/No (1-0)
13	understand gestures such as waving	Motion of hands or body to emphasize or help to express a thought or feeling
14	Epilepsy (L-1;M-2;H-3)	A disorder of the central nervous system characterized by loss of consciousness and convulsions (l-low; M-Medium; H-High).
15	Epileptic EEG abnormality (in 7 CDD and 82 ADSL)	-
16	Mutism (SL-1;SL-2;SL-3)	The condition of being unable or unwilling to speak
17	Having a first-degree relative with a developmental disorder	-
18	Severe intellectual disability (IQ < 35)	Intensely or extremely bad or unpleasant in degree or quality
19	Childhood Autism Rating Scale-Tokyo Version (L=1;M=2;H=3)	-
20	IQ Total score	-
21	15 items (score ≥ 2)(Y-1;N-0)	Yes/No (1-0)
22	1. Relationships with people	Yes/No (1-0)
23	2. Imitation	Copying (or trying to copy) the actions of someone else Yes/No (1-0)
24	3. Emotion	Yes/No (1-0)
25	4. Use of body (stereotypy)	Yes/No (1-0)
26	5. Relation to nonhuman objects (inappropriate interest in or use of objects)	not belonging to or produced by or appropriate to human beings Yes/No (1-0)
27	6. Adaptation to environmental change	Yes/No (1-0)
28	7. Visual responsiveness	Yes/No (1-0)
29	8. Auditory responsiveness	Yes/No (1-0)
30	9. Near receptor responsiveness	Yes/No (1-0)
31	10. Anxiety	A relatively permanent state of worry and nervousness occurring in a variety of mental disorders, usually accompanied by compulsive behavior or attacks of panic. Yes/No (1-0)
32	11. Verbal communication	Yes/No (1-0)
33	12. Nonverbal communication	Yes/No (1-0)
34	13. Activity level	(l-low; M-Medium; H-High).
35	14. Intellectual functioning (uneven intellectual profile)	Of or associated with or requiring the use of the mind. Yes/No (1-0)
36	15. General impressions	Yes/No (1-0)
37	Routines or repetitive behaviors	Yes/No (1-0)
38	repeating words	Yes/No (1-0)
39	repeating actions over and over	Yes/No (1-0)
40	obsessively following routines	Yes/No (1-0)
41	schedules for their actions	Yes/No (1-0)
42	Want to have routines where things stay the same so they know what to expect.	Yes/No (1-0)
43	Increase sensorimotor skills	relating to the sensory and motor coordination of an organism or to the controlling nerves .Yes/No (1-0)

44	Are behaviors linked with particular object?	Yes/No (1-0)
45	Do they occur in response to the same situation?	Yes/No (1-0)
46	Do they occur in a predictable pattern?	-
47	What activity or event preceded the behavior?	-
48	What environmental factors impact behavior?	-
49	What was the child reacting to?	-
50	Warning signs or predictors of the behavior	-
51	Restlessness	-
52	Eye aversion	A feeling of intense dislike. Yes/No (1-0)
53	Distractibility	Capable of being drawn aside or distracted
54	Pause	Interrupt temporarily an activity before continuing
55	Louder voice	Yes/No (1-0)
56	Hand flapping	Yes/No (1-0)
57	Vision (sight) - Sense - Over-Sensitive	Yes/No (1-0)
58	Vision (sight) - Sense - Under-Sensitive	Yes/No (1-0)
59	Vision (Sound) - Sense - Over-Sensitive	Yes/No (1-0)
60	Vision (Sound) - Sense - Under-Sensitive	Yes/No (1-0)
61	Vision (Smell) - Sense - Over-Sensitive	Yes/No (1-0)
62	Vision (Smell) - Sense - Under-Sensitive	Yes/No (1-0)
63	Vision (Touch (Tactile)) - Sense - Over-Sensitive	Yes/No (1-0)
64	Vision (Touch (Tactile)) - Sense - Under-Sensitive	Yes/No (1-0)
65	Vision (Taste) - Sense - Over-Sensitive	Yes/No (1-0)
66	Vision (Taste) - Sense - Under-Sensitive	Yes/No (1-0)
67	Vision (Movement) - Sense - Over-Sensitive	Yes/No (1-0)
68	Vision (Movement) - Sense - Under-Sensitive	Yes/No (1-0)
69	Low Tech/No Tech Adaptations for the Visual Environment - -Reduce bright lighting	Yes/No (1-0)
70	Low Tech/No Tech Adaptations for the Visual Environment-Decrease clutter	Yes/No (1-0)
71	Low Tech/No Tech Adaptations for the Auditory Environment	Yes/No (1-0)
72	Low Tech/No Tech Adaptations for the Vestibular (movement/balance) Environment	Yes/No (1-0)
73	Low Tech/No Tech Adaptations for the Proprioceptive (body position) Environment	Yes/No (1-0)
74	Change routine	Yes/No (1-0)
75	child does not move independently	Yes/No (1-0)

V. RESULT AND DISCUSSION

The feature selection algorithm is used to reduce the size of the feature subset. The optimization of the feature selection is to enhance the classification accuracy in the classification of ASD children in the e-learning process. The proposed relevant feature selection method has evaluated by the following metrics. Classification Accuracy, Kappa Statistics, True Positive Rate (TPR), False Positive Rate (FPR), Precision, Recall, Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Root Absolute Error (RAE), Root Relative Absolute Error (RRAE), and Receiver of Characteristics (ROC). Table 2, table 3, table 4 and table 5 depicts the result obtained by using Information Gain feature selection, Chi-Square feature selection, Gain Ratio Feature selection and ReliefF feature selection methodology. From these tables, the feature selection methods like Information Gain and Chi-Square gives the less number of features when it is compared with other two methods.

TABLE 2: RESULT OF THE INFORMATION GAIN FEATURE SELECTION METHOD ON ASD DATASET

Feature Number	Rank	Feature Name
1	0.654	Severe intellectual disability (IQ < 35)
2	0.3711	Total score ? 30
3	0.321	verbal (1)/non-verbal communication(0)
4	0.2663	Childhood Autism Rating Scale-Tokyo Version (L=1;M=2;H=3)
5	0.2635	repeating words

6	0.196	Lack of empathy(L-1:M-2:H-3)
7	0.1943	Fearfulness
8	0.1943	At first visit (Fear-1;Less Fear-0)
9	0.1778	interact with other people
10	0.1491	understand gestures such as waving
11	0.1487	adaptive behaviors
12	0.1472	Increase sensorimotor skills
13	0.1472	Pause
14	0.1472	Eye aversion
15	0.1472	Vision (sight) - Sense - Under-Sensitive
16	0.1472	Vision (Touch (Tactile)) - Sense - Under-Sensitive
17	0.1472	Low Tech/No Tech Adaptations for the Visual Environment-Decrease clutter
18	0.1472	Warning signs or predictors of the behavior
19	0.1393	schedules for their actions
20	0.1393	Louder voice
21	0.1364	Stereotyped behaviors
22	0.123	child does not move independently
23	0.1161	Head-growth decelerations
24	0.1142	Environmental factors impact behavior?
25	0.1065	Do they occur in a predictable pattern?
26	0.1065	Are behaviors linked?
27	0.1065	Do they occur in response to the same situation?
28	0.1065	Low Tech/No Tech Adaptations for the Auditory Environment
29	0.1065	Low Tech/No Tech Adaptations for the Proprioceptive (body position) Environment
30	0.1065	Low Tech/No Tech Adaptations for the Vestibular (movement/balance) Environment
31	0.1065	Vision (Movement) - Sense - Over-Sensitive
32	0.1065	Vision (Taste) - Sense - Over-Sensitive
33	0.1065	Vision (Taste) - Sense - Under-Sensitive
34	0.1065	Hand flapping
35	0.1065	Distractibility
36	0.1065	Vision (Smell) - Sense - Over-Sensitive
37	0.1065	Vision (Sound) - Sense - Over-Sensitive
38	0.1065	Vision (Sound) - Sense - Under-Sensitive
39	0.1012	Change routine
40	0.1012	What activity or event preceded the behavior?
41	0.1012	Vision (Smell) - Sense - Under-Sensitive
41	0.1012	Vision (Movement) - Sense - Under-Sensitive
42	0.1012	repeating actions over and over
43	0.0919	obsessively following routines
44	0.0814	Vision (sight) - Sense - Over-Sensitive
45	0.0814	Vision (Touch (Tactile)) - Sense - Over-Sensitive
46	0.0814	Low Tech/No Tech Adaptations for the Visual Environment - -Reduce bright lighting
47	0.0814	Routines or repetitive behaviors
48	0.0814	want to have routines where things stay the same so they know what to expect.

TABLE 3: RESULT OF THE CHI-SQUARED FEATURE SELECTION METHOD ON ASD DATASET

Feature Number	**Rank**	**Feature Name**
1	743.851	Severe intellectual disability (IQ < 35)
2	718.4872	Total score ? 30
3	274	verbal (1)/non-verbal communication(0)
4	204.8895	repeating words
5	192.6558	Childhood Autism Rating Scale-Tokyo Version (L=1;M=2;H=3)
6	185.0504	At first visit (Fear-1;Less Fear-0)
7	185.0504	Fearfulness
8	165.7819	adaptive behaviors
9	134.4529	stereotyped behaviors
10	107.1172	interact with other people
11	104.3431	Lack of empathy(L-1: M-2: H-3)
12	100.7233	Vision (sight) - Sense - Under-Sensitive
13	100.7233	Increase sensorimotor skills
14	100.7233	Warning signs or predictors of the behavior
15	100.7233	Pause
16	100.7233	Eye aversion
17	100.7233	Vision (Touch (Tactile)) - Sense - Under-Sensitive
18	100.7233	Low Tech/No Tech Adaptations for the Visual Environment-Decrease clutter
19	87.7853	the child does not move independently
20	58.8175	Change routine
21	58.8175	Vision (Movement) - Sense - Under-Sensitive
22	58.8175	Vision (Smell) - Sense - Under-Sensitive
23	58.8175	What activity or event preceded the behavior?
24	58.8175	repeating actions over and over
25	42.4051	understand gestures such as waving

26	38.6492	Louder voice
27	38.6492	schedules for their actions
28	30.9162	Environmental factors impact behavior?
29	30.7736	Head-growth decelerations
30	28.4311	Are behaviors linked?
31	28.4311	Vision (Taste) - Sense - Under-Sensitive
32	28.4311	Vision (Smell) - Sense - Over-Sensitive
33	28.4311	Vision (Taste) - Sense - Over-Sensitive
34	28.4311	Low Tech/No Tech Adaptations for the Vestibular (movement/balance) Environment
35	28.4311	Low Tech/No Tech Adaptations for the Proprioceptive (body position) Environment
36	28.4311	Vision (Movement) - Sense - Over-Sensitive
37	28.4311	Low Tech/No Tech Adaptations for the Auditory Environment
38	28.4311	Distractibility
39	28.4311	Do they occur in response to the same situation?
40	28.4311	Do they occur in a predictable pattern?
41	28.4311	Vision (Sound) - Sense - Over-Sensitive
41	28.4311	Vision (Sound) - Sense - Under-Sensitive
42	28.4311	Hand flapping
43	25.0876	obsessively following routines
44	21.3936	Vision (sight) - Sense - Over-Sensitive
45	21.3936	Vision (Touch (Tactile)) - Sense - Over-Sensitive
46	0.0814	Low Tech/No Tech Adaptations for the Visual Environment - -Reduce bright lighting
47	0.0814	Routines or repetitive behaviors
48	0.0814	want to have routines where things stay the same so they know what to expect.

TABLE 4: RESULT OF THE GAIN RATIO FEATURE SELECTION METHOD ON ASD DATASET

Feature Number	Rank	Feature Name
1	1	verbal (1)/non-verbal communication(0)
2	0.6753	repeating words
3	0.6053	At first visit (Fear-1;Less Fear-0)
4	0.6053	Fearfulness
5	0.59	Total score ? 30
6	0.54	adaptive behaviors
7	0.4688	stereotyped behaviors
8	0.3491	interact with other people
9	0.3252	Pause
10	0.3252	Warning signs or predictors of the behavior
11	0.3252	Eye aversion
12	0.3252	Vision (sight) - Sense - Under-Sensitive
13	0.3252	Increase sensorimotor skills
14	0.3252	Low Tech/No Tech Adaptations for the Visual Environment-Decrease clutter
15	0.3252	Vision (Touch (Tactile)) - Sense - Under-Sensitive
16	0.287	child does not move independently
17	0.2854	Severe intellectual disability (IQ < 35)
18	0.2199	Childhood Autism Rating Scale-Tokyo Version (L=1;M=2;H=3)
19	0.2038	Lack of empathy(L-1:M-2:H-3
20	0.203	What activity or event preceded the behavior?
21	0.203	Vision (Smell) - Sense - Under-Sensitive
22	0.203	Vision (Movement) - Sense - Under-Sensitive
23	0.203	Change routine
24	0.203	repeating actions over and over
25	0.1493	understand gestures such as waving
26	0.1394	Louder voice
27	0.1394	schedules for their actions
28	0.1272	Environmental factors impact behavior?
29	0.1187	C_Gender
30	0.1187	Head-growth decelerations
31	0.1125	Are behaviors linked?
32	0.1125	Vision (Taste) - Sense - Under-Sensitive
33	0.1125	Vision (Smell) - Sense - Over-Sensitive
34	0.1125	Vision (Taste) - Sense - Over-Sensitive
35	0.1125	Low Tech/No Tech Adaptations for the Vestibular (movement/balance) Environment
36	0.1125	Low Tech/No Tech Adaptations for the Proprioceptive (body position) Environment
37	0.1125	Vision (Movement) - Sense - Over-Sensitive
38	0.1125	Low Tech/No Tech Adaptations for the Auditory Environment
39	0.1125	Distractibility
40	0.1125	Do they occur in response to the same situation?
41	0.1125	Do they occur in a predictable pattern?
42	0.1125	Vision (Sound) - Sense - Over-Sensitive
43	0.1125	Vision (Sound) - Sense - Under-Sensitive
44	0.1125	Hand flapping
45	0.0925	obsessively following routines
46	0.0828	Vision (sight) - Sense - Over-Sensitive
47	0.0828	Vision (Touch (Tactile)) - Sense - Over-Sensitive

48	0.0828	Low Tech/No Tech Adaptations for the Visual Environment - -Reduce bright lighting
49	0.0828	Routines or repetitive behaviors
50	0.0828	want to have routines where things stay the same so they know what to expect.

TABLE 5: RESULT OF THE RELIEFF FEATURE SELECTION METHOD ON ASD DATASET

Feature Number	Rank	Feature Name
1	0.697	Severe intellectual disability (IQ < 35)
2	0.496	Childhood Autism Rating Scale-Tokyo Version (L=1;M=2;H=3)
3	0.474	understand gestures such as waving
4	0.428	Louder voice
5	0.428	schedules for their actions
6	0.405	C_Gender
7	0.405	Head-growth decelerations
8	0.403	obsessively following routines
9	0.397	Vision (sight) - Sense - Over-Sensitive
10	0.397	Vision (Touch (Tactile)) - Sense - Over-Sensitive
11	0.397	want to have routines where things stay the same so they know what to expect
12	0.397	Routines or repetitive behaviors
13	0.397	Low Tech/No Tech Adaptations for the Visual Environment - -Reduce bright lighting
14	0.39	Lack of empathy(L-1:M-2:H-3)
15	0.36	verbal (1)/non-verbal communication(0)
16	0.357	Distractibility
17	0.357	Low Tech/No Tech Adaptations for the Proprioceptive (body position) Environment
18	0.357	Do they occur in response to the same situation?
19	0.357	Are behaviors linked?
20	0.357	Vision (Movement) - Sense - Over-Sensitive
21	0.357	Do they occur in a predictable pattern?
22	0.357	Low Tech/No Tech Adaptations for the Vestibular (movement/balance) Environment
23	0.357	Vision (Smell) - Sense - Over-Sensitive
24	0.357	Vision (Sound) - Sense - Under-Sensitive
25	0.357	Vision (Sound) - Sense - Over-Sensitive
26	0.357	Low Tech/No Tech Adaptations for the Auditory Environment
27	0.357	Vision (Taste) - Sense - Under-Sensitive
28	0.357	Vision (Taste) - Sense - Over-Sensitive
29	0.357	Hand flapping
30	0.338	interact with other people
31	0.338	Environmental factors impact behavior?
32	0.337	repeating words
33	0.329	Total score ? 30
34	0.226	Fearfulness
35	0.226	At first visit (Fear-1;Less Fear-0)
36	0.222	Vision (Smell) - Sense - Under-Sensitive
37	0.222	Vision (Movement) - Sense - Under-Sensitive
38	0.222	What activity or event preceded the behavior?
39	0.222	repeating actions over and over
40	0.222	Change routine
41	0.219	Warning signs or predictors of the behavior
42	0.219	Vision (Touch (Tactile)) - Sense - Under-Sensitive
43	0.219	Increase sensorimotor skills
44	0.219	Vision (sight) - Sense - Under-Sensitive
45	0.219	Eye aversion
46	0.219	Pause
47	0.219	Low Tech/No Tech Adaptations for the Visual Environment-Decrease clutter
48	0.218	child does not move independently
49	0.187	stereotyped behaviors
50	0.183	adaptive behaviors

TABLE 6: PERFORMANCE EVALUATION OF THE ORIGINAL DATASET, INFORMATION GAIN, CHI-SQUARE, GAIN RATIO AND RELIEF FEATURE SELECTION METHODS

Evaluation Metrics	Original Dataset	Information Gain Feature Selection	Chi-Square Feature Selection	Gain Ratio Feature Selection	ReliefF Feature Selection
Classification Accuracy	71.92	77.49	75.48	74.35	73.45
Kappa Statistic	0.26	0.51	0.55	0.45	0.42
Mean absolute error	0.35	0.27	0.25	0.30	0.32
Root mean squared error	0.49	0.43	0.39	0.50	0.51
Relative absolute error	64.78	55.26	49.18	63.24	60.06
Root relative squared error	88.67	86.58	75.66	87.45	88.02
True Positive Rate	0.68	0.70	0.79	0.71	0.65
False Positive Rate	0.37	0.35	0.27	0.36	0.37
Precision	0.68	0.70	0.76	0.71	0.73
Recall	0.68	0.70	0.76	0.72	0.74
ROC Area	0.39	0.73	0.85	0.71	0.72

From the above table 6, the metrics like Accuracy, Precision, Kappa Statistics, TPR achieves higher value for Information Gain and Chi-square when the classification technique Artificial Neural Network is utilized. The other error rates are also less for these two methods.

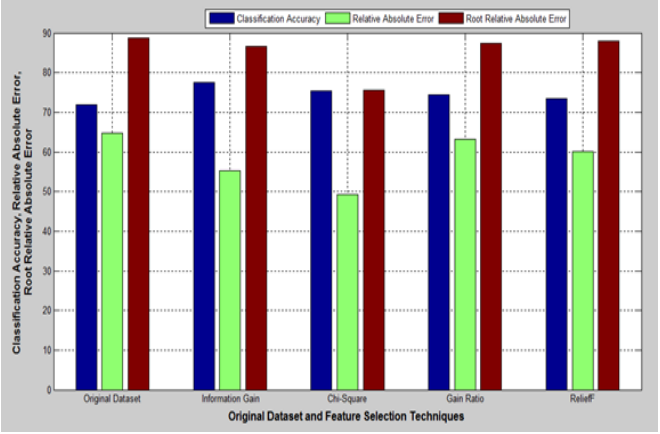

Figure 1: Performance Analysis of feature selection techniques on Classification Accuracy, Information Gain, Chi-Square, Gain Ratio and ReliefF using ANN classification

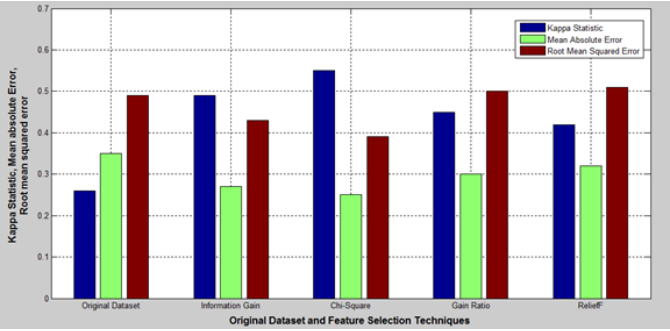

Figure 2: Performance analysis of Information Gain, Chi-Square, Gain Ratio and ReliefF feature selection method on Kappa Statistic, Mean Absolute Error, Root Mean Squared Error using ANN classification

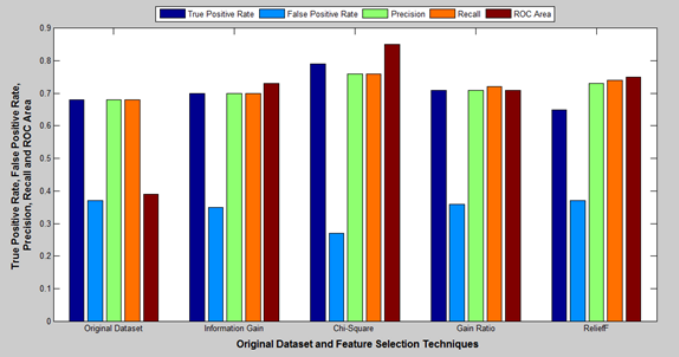

Figure 3: Performance analysis of Original dataset, Information Gain, Chi-Square, Gain Ratio, ReliefF feature selection techniques on TPR, FPR, Precision, Recall, ROC area using ANN classification

The above figures 1, 2 and 3 depicts the performance analysis of the original dataset (without pre-processing), Information Gain, Chi-Square, Gain Ratio, ReliefF feature selection techniques on Classification accuracy, Kappa Statistic, Relative absolute error, root relative absolute error, Kappa Statistic, Mean Absolute Error, Root Mean Squared Error, TPR, FPR, Precision, Recall, ROC area..

.

VI. CONCLUSION

In this paper, the feature selection methodologies are compared to reduce the size of the feature set. The problems are dealing with the dataset are having a lot of features, few samples and eventually, suffering from imbalanced distribution of dataset. From the above results, it is concluded that the Information Gain and Chi-Square feature selection methods performed well than the other existing methods.

REFERENCES

[1] Eman Ahmed Zaky, "Autism Spectrum Disorder (ASD); The Past, The Present, and The Future", *Journal of Child & Adolescent Behavior*, pp.1-4, 2017.

[2] Bappaditya Adak and Santoshi Halder, "Systematic Review on Prevalence for Autism Spectrum Disorder with Respect to Gender and Socio-Economic Status", *Journal of Mental Disorders J and Treatment*, Volume 3, Issue 1, pp.1-9, 2017.

[3] Eric Fombonne, "Epidemiology of Pervasive Developmental Disorders", *Pediatric Research*, Vol. 65, No. 6, pp. 591-598, 2009.

[4] Milos Radovic, Mohamed Ghalwash, Nenad Filipovic, and Zoran Obradovic, "Minimum redundancy maximum relevance feature selection approach for temporal gene expression data", *BMC Bioinformatics*, 18:9, pp.1-14, 2017.

[5] Pream Sudha V, "Feature Selection Techniques for the Classification of Leaf Diseases in Turmeric", *International Journal of Computer Trends and Technology (IJCTT)*, Volume 43 Number 3, pp.138-142, January 2017.

[6] Pream Sudha V, "Feature Selection Techniques for the Classification of Leaf Diseases in Turmeric", *International Journal of Computer Trends and Technology (IJCTT)*, Volume 43 Number 3, pp.138-142, January 2017.

[7] Omer Faruk ERTUGRUL, Mehmet Emin TAGLUK, "A fast feature selection approach based on extreme learning machine and coefficient of variation", *Turkish Journal of Electrical Engineering & Computer Sciences*, pp. 3409 – 3420, 2017.

[8] B. S. Harish, M. B. Revanasiddappa, "A Comprehensive Survey on various Feature Selection Methods to Categorize Text Documents", *International Journal of Computer Applications*, Volume 164 - No.8, pp.1-7, 2017.

[9] Szidonia Lefkovtis, Laszlo Lefkovits, "Gabor Feature Selection based on Information Gain", *10th International Conference Interdisciplinary in Engineering*, pp. 892-898, 2017.

[10] R.P.L. Durgabai, "Feature Selection using ReliefF Algorithm", *International Journal of Advanced Research in Computer and Communication Engineering*, Vol. 3, Issue 10, pp. 8215-8218, October 2014.

