Predicting Autism Spectrum Disorder in Infants Using Machine Learning

Ms. Rakhee Kundu¹, Mr. Suranjan Das²

¹Assistant Professor Computer Department- Mukesh Patel School Of Technology Management and Engineering, NMIMS University, Mumbai, India
²Sr. Technical Associate, Tech Mahindra

Abstract—Machine Learning techniques have played an important role in identifying patterns, grouping similar objects together and modelling data. These techniques learn from the given data and are capable of making prediction of outcomes for new datasets. In applications where the data is too large and cannot be manually analysed, applying machine learning techniques has helped in automation of many process. This has but, not been done often and hardly ever been done for ASD and developing a decision support system that helps physicians with the detection of ASD, has scarcely been done, that is tried by the shortage of literature on the utilization of it.

Keywords—Machine Learning, Autism Spectrum Disorder, Classification, Decision tree, Regression

1. INTRODUCTION

The children of today who are going to be the future of this nation are dealing with hurdles in their mental development like ASD also known as Autism Spectrum Disorder. Autism is not hard to detect but it requires a lot of learning and training for physicians to detect it. Dr. Mark Hyman says that “Autism is not a genetic brain disorder but a systematic body disorder that affects the brain”. A toxicant setting triggers sure genes in individuals prone to this condition. ASD are the neurological disorder that creates a lifelong disorder in children resulting in mental illness. Research says that there are more than 1 million cases in India. Treatment can help but it can’t be cured. Currently ASD is detected by understanding the behavioural and intellectual activities of a child. This diagnosis can be subjective, time consuming, inconclusive, does not provide proper insight of his genetics and is not suitable for early detection.

I. WHAT IS AUTISM

The autism spectrum disorder belong to an umbrella class category of five childhood onset conditions called Pervasive Developmental Disorder (PDD). There are three most common PDD.

1. Asperger Syndrome: People suffering from this disorder have milder symptoms of autistic disorder and have social challenges and unusual behaviours. Although they are doing not suffer from language or intellectual incapacity.

2. Autistic Disorder: Also known as classic Autism. People stricken by this disorder have vital language delays, social and communication challenges, unusual behaviours and interest and may also have intellectual disability. Most of the common behaviour shown by children suffering from Autism are:

- Aggression (hitting, kicking, scratching etc.)
- Self Injury (head-banging, hand-biting, hitting walls etc.)
- Disruption (interrupting, yelling, knocking things over etc.)
- Elopement (wandering, bolting etc.)
- Stereotype (hand-flapping, toe-walking, rocking etc.)
- Tantrums (crying, screaming, defiant behaviour etc.)
- Non-compliance (disobeying directions, whining etc.)
- Obsession (repeatedly talking about the same topic, perseverance etc.)

3. Pervasive Developmental Disorder- Not Otherwise Specified: Also known as a typical Autism. The individuals that meet just some of the symptoms of Autism or Asperger Syndrome are classified underneath this class. Childhood disintegrative and Rett Syndrome are the other pervasive developmental disorder. As results of each are extremely rare genetic diseases they are sometimes thought of to be separate medical conditions that do not really belong on the autism.

2. LITERATURE REVIEW

There is no specific treatment to cure autism but some preference using robotic toys have proved as a help for detecting social interaction to children with ASD which are as follows:

1. Michael Siller and Marian Sigman proposed developmental relationships connecting the child’s subsequent development of communication skills and parental understanding in adult with ASD [1].
2. Tony Charman et al. described an consideration should be directed at assessing these skills in 2 and 3 year old children referred for a diagnosis of autism spectrum disorder [2].
3. G. Leroy et al. proposed a data mining techniques for decision tree and association rule using to predict autism affect child behavior provided more detailed insights for high and low levels of appropriate and inappropriate behavior [3].
4. S.Wheelwright et al. investigating by the relationship between scores on the Empathy Quotient, Systemizing Quotient-Revised and Autism Spectrum Quotient in both a large sample of typical participants and sample of adults with autism spectrum conditions (ASC) [4].
5. Sheena angra and Sachin ahuja described an predict grades of students of one class by analyzing the grades of previous classes using K-Means, Linear Regression and Decision Tree in Rapid Miner environment [5].
6. Brian P. Keane et al. suggest that the ability to integrate between the auditory and visual sense modalities is unimpaired among high-functioning adults with autism [6].
7. A.Martin et al. used to the hybrid model of genetic algorithm, Fuzzy c-means algorithm, MARS for prediction of bankruptcy [7].
8. Tanaya Guha et al. describe reduced complexity in facial expression dynamics of subjects with High Functioning Autism relative to their typically developing peers. Significant difference is observed for expressions related to disgust, joy and sadness [8].
9. Ahmed Hassan et al. used Naïve Bayes (NB), Random Forest (RF), Decision Trees (C4.5), Support Vector Machines (SVM), K-Nearest Neighbor (KNN) and Logistic Regression (LR) to evaluate the classification performance by using three microarray gene expression datasets namely the Leukemia Cancer, Lung cancer and breast Cancer datasets [9].
10. Niyati Gupta et al. used to the various classification methods and compared the results of various algorithms on Weka on the basis of accuracy, sensitivity and specificity with the results of various other algorithms implemented on Matlab and Rapid Miner [10].
11. E. M. Albornoz et al. presented an evolutionary method for automatic selection of features of speech in a classification task for autism and It is based on a genetic algorithm that selects the best combination of acoustic features using support vector machines as classifier [11].
12. M.S. Mythili and A.R.Mohamed Shanavas comparisons on data mining classification methods used to predict learning skills of autistic children. The classification algorithms like, J48 and Support Vector Machine (SVM) algorithms are compared check the efficiency and outperform the SVM [12].
13. Vitthal Manekar and Kalyani Waghmare comparisons on various data mining algorithms like, decision tree, Naïve Bayes, KNN and SVM and Hybrid Approach of SVM algorithm and the SVM algorithm with evolutionary approach is more Accurate [13].

14. Parvathi I and Siddharth Rautaray describes about the proposal of hybrid data mining model to extract classification knowledge for various diseases in clinical decision system and presents a framework of the tool various tools used for analysis [14].

15. Priyanka Juneja and Anshul Anand has been analysis on the ASD patient symptom based a pre-level decision is taken about to recognize the probability of ASD [15].

16. Mohana E and Poonkuzhali.S describes the result for best suited classifier using to predicting the risk level of ASD [16].

17. Motaz M. H. Khorsheid et al. proposed an hybrid machine learning models using to predict terrorist groups responsible of attacks in Middle East and North Africa [18].

18. Ionuț Taranu describes data mining techniques is most beneficial in health care organization [19].

19. Priyanka Sanjay Podutwar and Prof. Ms. R. R. Tuteja proposed a soft computing technique using to find the accuracy of ASD [20].

20. Ashmeet Singh and R Sathyaraj describe various Methodologies used to predicting small and large dataset in rapid miner environment [21].

21. Aggarwal and Zhai propose that these are the most popular algorithms for text classification, they also state they are just as capable to classify the structured data. The classification algorithms to be used are: Naive Bayes [17], Random Forest[22] and Support Vector Machine[23]. Besides these three popular algorithms, they have also used an older method J48 [24], which is a java implementation of a well known algorithm called C4.5[25], to be able to compare the Random Forest method with a different tree algorithm.

22. Rapid Miner combines tools and applicability to provide user-friendly integration environment of the newest data mining techniques [26].

23. Rapid Miner is open-source data mining with java software and platform for data science software. It is provides an integrated environment for data preparation, deep learning, machine learning, predictive analysis and text mining. Rapid Miner is developed on open core model. An application of Rapid Miner covers a wide range of real-world data mining task. Rapid Miner tool supports the DRN hybrid algorithm. [27].

3. RESEARCH QUESTIONS OR HYPOTHESIS

Research Questions or Hypothesis

How can machine learning be used to detect Autism Spectrum Disorder?

This question may be divided within the following analysis queries

1. Which conditions co-occur with Autism Spectrum Disorder?

2. Does the data contain Co-occurring conditions with Autism Spectrum Disorder?

3. Are the Co-occurring conditions prognostic of Autism Spectrum Disorder using Machine Learning?

4. PROPOSED MODEL

Our proposed model will follow mining information from both structured and semi-structured data. Our approach to build and train the models will follow two ways.

1. The first way to collect data we will use a gaming app in which they can copy the scenes as shown in the video. The app can be used by the parents or authorized persons using smart phone mobile app. Then we will capture the specific behaviour of the child in response to some videos in the absence of audio by the inbuilt camera of the app which will be recorded in some given time.
frame. This will be called as Test Exercises which can be conducted by the parents at their home or by authorized persons at their clinical environment.

2. The second way to collect the data will follow a structured way which will be a health questionnaire to be filled up by the parents to understand the genetic history or background history of the child.

Both this data will be given to a single assessment tool where it will be trained to identify whether a child is a ASD or Not ASD. The targeted age group to test ASD will be from 2 to 6 years. We will propose a robust machine learning approach to tackle a challenging problem that involves mining from semi-structured and structured in video and audio format. We will collect the data from the structured format from the parental questionnaire and semi-structured video screening methodology done by parents at home or at clinical environment combining them together into a single assessment tool. We will apply different word embedding models including the state-of-art distributed representations for semi-structured audio and video data and establish a promising baseline for automated ASD detection on such a dataset.

5. FEATURE DETECTION

A feature in machine learning is any column on input matrix that we are using as an independent variable. Feature can be taken from source data but it would be better if we do some kind of transformation so that we get the raw data into its most appropriate form for the model. We will collect the data on some features which will provide a list of Co-occurring conditions such as developmental delay, obesity, less physical activity, premature birth, social interaction delay, learning disability etc. These co-occurring conditions will then be accustomed to facilitate us to presume ASD.

The most popular technique used till now to evaluate ASD is the brain MRI and EEG signal. In brain MRI they could observe brain overgrowth which is a biomarker for ASD. They have seen the children with ASD the cortical surface area of the brain grew faster as compared with infants who do not have ASD. EEG signals are used to measure the brain’s electrical activity that can accurately predict or rule out ASD in infants. It is used in many paediatric settings.

The machine learning algorithms for detection of ASD used till now are described as follows:

1. Naïve Baye’s Algorithm: It is supervised machine learning algorithm. By the term Naïve we mean immature. A Naïve Baye’s classifier is a simple probability based algorithm. It uses Baye’s theorem. It is often associated with NLP (Natural Language Processing) applications like spam recognition or sentiment analysis. It assumes instances are independent of each other. They work very well in complex real world scenarios. It offers a fast building and scoring both for binary and multiclass situations for relatively low volume of data.

   Advantages:
   1. The algorithm as well as it’s implementation process is simple. Calculations do not involve any numerical optimization, no matrix algebra and no calculus involved.
   2. The algorithm is efficient to train and use. It can be easily updated with new data. This algorithm is similar to linear classifier and hence is fast too.
   3. From one dataset we can estimate content features from message without headers.
   4. Unlabelled data can be used in parameter estimation from other applications. This algorithm performs well with smaller training sets.

   Disadvantages:
   1. The independence assumption is a very strong inference and untrue for most real world problem.
   2. The boundaries between classes have to be fine tuned and not to be set analytically.
3. it is not effective algorithm as it is simple.

2. **ADA Boost:** In real life when we are trying to make an important decision we are taking advice of multiple persons instead of trusting one single person. The same logic is applied to machine learning using meta algorithm. Meta algorithms are a style of bringing together other algorithms. The most popular meta algorithm is ADA Boost. This is a powerful tool as it is considered to be the best supervised learning algorithm. ADA Boost algorithm is for boosting. In this we build a decision stump classifier which is a single node decision tree. The ADA boost algorithm is then applied to decision stump classifier to work on difficult dataset and see how quickly outperforms other classification methods. It works with numeric and nominal values.

   **Advantages:**
   1. Low generalization error.
   2. Easy to use.
   3. Works with the most classifiers.
   4. No parameters to adjust.

   **Disadvantages:** Sensitive to outliers.

3. **Bagging:** Bootstrap aggregating which is also known as bagging is a technique where the data is taken from the original dataset to make a new datasets. The datasets are of the same size as the original. Each dataset is built randomly by selecting an example from the original with replacement. By “with replacement” I mean you can select the same example more than once. This property allows you to have values in the new dataset more that are repeated and some values from the original won’t be present in the new set. After datasets are built a learning algorithm is applied to each one individually. When we would like to classify a new piece of data that would take a majority vote. There are more advanced methods of bagging such as Random forest.

4. **Decision Tree:** In decision trees we break the classifier down into a set of choices about each feature starting at the root of the tree band progressing down to leaves where we receive the final classification decision. The data structure is very easy to understand and implement. In this approach we can turn into a set of logical disjunction by applying if-then-else which are useful for use in production system. In this each node in the tree specifies a test of some attribute of the instance and each branch descending from that node corresponds to one of the possible values for this attribute specified by this by this node then moving down the tree branch corresponding to the value of the attribute in the given example. This method is continued for the sub-tree rooted at the new node.

   **Disadvantages:**
   1. Instances are represented by attribute value pairs.
   2. Target function has discrete output values.
   3. Disjunctive disruptions are required.
   4. Training data may contain errors,
   5. Training data may contain missing attribute values.

5. **Random Forest:** Random forest is a supervised flexible easy to use machine learning algorithm that produces even without hyper-parameter tuning a great result most of the time. It is one of the most used algorithms because of it’s simplicity and it can be used both for classification and regression tasks. It creates a forest which is an ensemble of decision tree with most of the time trained with the bagging method. Instead of searching the most important feature while splitting the node it searches for the best feature among a random subset of features. Another important quality of random forest is it measures the relative importance of each feature on the prediction.
6. **Support Vector Machine (SVM):** SVM is a statistical based learning algorithm that has been widely used by researchers in various fields like businesses, text categorization, pattern recognition and protein function prediction. SVM belongs to the type of supervised learning methods used for classification and regression. An SVM model is representation of the example points in space mapped so that the example of the separate are divided by a clear gap that is as wide as possible. It is often used in industrial applications either when data are not labelled or when only some of the data are labelled as a pre-processing for a classification pass. SVM can handle any classification, clustering, regression and even novelty detection problems.

**Advantages:**

1. They cannot only solve binary class pattern recognition problems but also for multiclass classification, regression estimation, feature selection etc.
2. Lesser chances of information loss as SVM kernel considers the dot product of the feature vectors in high dimensional space to construct the optimal hyper-plane rather than clustering or interpolating in input space. SVM can handle a large number of attributes of a dataset unlike neural nets as these traditional algorithms needed reduction of dimensionality i.e. the number of attributes before they can be applied to a high dimensional problem.
3. Even in the field of bioinformatics it has proven to be useful as there we have more than 20,000 attributes.
4. They were very versatile as they have many extensions too like support vector, decision tree, neural network based SVM, fuzzy SVM etc.
5. SVM can also perform a nonlinear classification using what is called the kernel trick i.e. implicitly mapping their inputs into high dimensional feature areas.

**Disadvantages:**

1. The best choice of kernel for a given problem is still an issue in research work.
2. It’s speed and size during training and testing is still an issue. Training for very large datasets/ millions of support vectors is still an unsolved problem.
3. It lacks efficiency in solving of multiclass problems. It is still an active research area.

6. **PROPOSED MODEL**

Firstly asking a sufficiently through set of questions to be able to effectively classify all the conditions would be prohibitive on any parent’s time. Secondly a large sample of training data would be needed of children with everyone of the conditions which condition features that are relevant not just to their particular condition but also to all other co-occuring conditions that the classifier attempt to identify. As shown in the diagram we propose a tree of classification algorithms. This tree would start with an algorithm to determine weather the child is at risk of any number of broad categories of conditions and if so, presents them with a new particular machine learning algorithm that is targeted to determine more precisely which kind of condition they have. Further classifiers can then be trained to identify a more specific condition if possible. The machine learning algorithms that run at each node can be individual algorithms or composite algorithms. The algorithms at each node of the tree can further come to an inconclusive determination at each node. The functionality to return inconclusive results is especially useful in tree of algorithms due to the fact that it allows for much more flexible tuning of the algorithms.

For e.g. if there is an algorithm that is unable to accurately classify 80% of children but which can determine a very accurate classification on the remaining 20% such an algorithm would be infeasible to use in a standalone screening application. However in a tree of algorithms the inconclusive parameter of the algorithm at that node can be tuned to reach to an inconclusive determination in the vast majority of subjects without losing relevance while giving a more precise determination for the minority of subjects for which an accurate conclusive determination is possible.
We will collect the data on some features which will also provide a list of co-occurring conditions such as

1. Developmental delay
2. Obesity
3. Less Physical activity
4. Premature birth
5. Very low birth weight
6. Social interaction delay
7. Learning disability
8. Speech or other language problems
9. Body mass index
10. Attendance to religious events
11. Vaccinations given
12. Allergic to any food products
13. Genetic history
14.

![Proposed Decision Tree Model to Predict ASD](image)

**Fig. 1: Proposed Decision Tree Model to Predict ASD**

### TABLE 1: STAGES OF RESEARCH

| Number | Task                  |
|--------|-----------------------|
| 1.     | **DEFINE**            |
| 1.1    | Define the criteria for Inclusion/Exclusion |
| 1.2    | Identify the fields of research |
| 1.3    | Determine the appropriate sources |
| 1.4    | Decide on the specific search terms |
| 2.     | **SEARCH**            |
| 2.1    | Search               |
| 3.     | **SELECT**           |
| 3.1    | Refine the sample    |
| 4.     | **ANALYZE**          |
| 4.1    | Open Coding          |
| 4.2    | Axial Coding         |
| 4.3    | Selective Coding     |
| 5.     | **PRESENT**          |
5.1 Represent and structure the content
5.2 Structure the Article

7. RESEARCH GAPS TO BE ADDRESSED

Discovering learning patterns at various time frames to study changes in progress and better prediction of the ASD for early intervention. Applying the brain connectivity measures for large population of ASD and typical children in order to effectively bring it in regular clinical practices. Creating a mobile application which allows parents or authorized people to avoid the child’s development. The database creation and training the models with very low loss of classifiers to determine ASD in children. To determine the co-occurring conditions along with ASD and to check if the co-occurring conditions with ASD predicted using machine learning techniques.

8. RESEARCH IMPLICATIONS AND FUTURE SCOPE

Early intervention of detection of ASD will help to reduce the impact of ASD in children and it may also be possible to completely abolish it from children with help of some therapies. With this techniques the physicians can detect the severity of ASD and help in child’s normal development.

Future scope of this research involves it can be predicted in advance before the birth of the child i.e. during pregnancy if a child is autistic or not. It is also seen that some autistic cases are detected after the vaccination of MMR (Measles, Mumps, Rubella). So physicians can determine the component of that vaccination that are stopping the developmental growth of a child and examine if those vaccinations can be given at the later stage. Researchers can also research will stem cell transfusion help for child’s development in Autism.

9. CONCLUSION

Machine learning methods have showed programming results in varying applications aside from ASD research. In particular recent research studies in machine learning claimed one or more of the following:

1. New ways to diagnose cases related to ASD
2. Shortening time associated with diagnosis process of ASD or at least self-administered autism trait tests.
3. Reducing the features associated with existing ASD tools without hindering sensitivity, specificity or the accuracy of the test.
4. Identifying the best ranked features that influence ASD.
5. Determining overlapped features among different types of attributes.
6. Determining the co-occurring conditions that are seen with ASD.

These techniques will help to identify the parents if a child is allergic to any food products that are stopping the developmental growth of children.

10. REFERENCES

[1]. Michael Siller and Marian Sigman, “The Behaviors of Parents of Children with Autism Predict the Subsequent Development of Their Children’s Communication”, Journal of Autism and Developmental Disorders, Vol. 32, No. 2, April 2002.

[2]. Tony Charman, Simon Baron-Cohen, John Swettenham, Gillian Baird, Auriol Drew and Antony Cox, “Predicting language outcome in infants with autism and pervasive
developmental disorder”, International Journal of Language & Communication Disorders, Vol.38, No.3, 2003.

[3]. G. Leroy, A. Irmscher, and M.H. Charlop-Christy, “Data Mining Techniques to Study Therapy Success with Autistic Children”, 2006 International Conference on Data Mining, 26 - 29 June 2006, Monte Carlo Resort, Las Vegas, USA.

[4]. S. Wheelwright, S. Baron-Cohen, N. Goldenfeld, J. Delaney, D. Fine, R. Smith, L. Weil and A. Wakabayashi, “Predicting Autism Spectrum Quotient (AQ) from the Systemizing Quotient-Revised (SQ-R) and Empathy Quotient (EQ)”, 2006.

[5]. Sheena angra and Sachin ahuja, “Analysis of student’s data using rapid miner”, Journal of Today’s Ideas – Tomorrow’s Technologies, Vol. 4, No. 1, June 2016 pp. 49–58.

[6]. Brian P. Keane, Orna Rosenthal, Nicole H. Chun and Ladan Shams, “Audiovisual integration in high functioning adults with autism”, Research in Autism Spectrum Disorder for Elsevier journal, Volume. 4, Issue 2, April-June 2010.

[7]. A.Martin, V.Gayathri, G.Saranya, P.Gayathri and Dr.Prasanna Venkatesan, “A Hybrid Model For Bankruptcy Prediction Using Genetic Algorithm, Fuzzy C-Means and Mars”, International Journal on Soft Computing ( IJSC ), Vol.2, No.1, February 2011.

[8]. Tanaya Guha, Member, IEEE, Zhaojun Yang, Student Member, IEEE, Ruth B. Grossman, and Shrikanth S. Narayan, Fellow, IEEE, “A Computational Study of Expressive Facial Dynamics in Children with Autism”, IEEE TRANSACTIONS ON AFFECTIVE COMPUTING, VOL. XX, NO. X, MARCH 2016.

[9]. Prof. Dr. Ahmed Hassan, Asistant Prof. Osama Abdo Mohamed, Prof. Dr. Ahmed Soufi Abou-Taleb, and Mr. Amr Hassan, “A Hybrid Feature Selection Approach Of Ensemble Multiple Filter Methods And Wrapper Method For Improving The Classification Accuracy Of Microarray Data Set”, IRACST - International Journal of Computer Science and Information Technology & Security (IJCITS), ISSN: 2249-9555 Vol. 3, No.2, April 2013.

[10]. Niyati Gupta, Arushi Rawal, Dr. V.L. Narasimhan, and Savita Shiwani, “Accuracy, Sensitivity and Specificity Measurement of Various Classification Techniques on Healthcare Data”, IOSR-Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661, p- ISSN: 2278-8727Volume 11, Issue 5 (May. - Jun. 2013), PP 70-73.

[11]. E. M. Albornoz, L. D. Vignolo, C. E. Martinez & D. H. Milone, "Genetic Wrapper Approach for Automatic Diagnosis of Speech Disorders related to Autism” 14th IEEE International Symposium on Computational Intelligence and Informatics (CINTI), nov, 2013.

[12]. M.S. Mythili and A.R.Mohamed Shanavas, “A Novel Approach to Predict the Learning Skills of Autistic Children using SVM and Decision Tree”, International Journal of Computer Science and Information Technologies, Vol. 5 (6), 2014, 7288-7291.

[13]. Viththal Manekar and Kalyani Waghmare, “ Improving Accuracy of SVM Using Hybrid Cultural Algorithm”, Int.J.Computer Technology & Applications,Vol 5 (3),1194-1197.

[14]. Parvathi I and Siddharth Rautaray, “Survey on Data Mining Techniques for the Diagnosis of Diseases in Medical Domain”, International Journal of Computer Science and Information Technologies, Vol. 5 (1), 2014, 838-846.

[15]. Priyanka Juneja and Anshul Anand, “Analyses of Autistic Patients By using Interpretation Value Analysis”, International Journal of Computer Science and Mobile Computing, Vol.3 Issue.7, July- 2014, pg. 585-593.

[16]. Mohana E and Poonkuzhali.S, “Categorizing The Risk Level Of Autistic Children Using Data Mining Techniques”, International Journal of Advance Research In Science And Engineering IJARSE, Vol. No.4, Special Issue (01), April 2015.

[17]. I. Kononenko. Inductive and Bayesian Learning in Medical Diagnosis. 1993.
[18]. Motaz M. H. Khorshid, Tarek H. M. Abou-El-Enien and Ghada M. A. Soliman, “Hybrid Classification Algorithms For Terrorism Prediction In Middle East And North Africa”, International Journal of Emerging Trends & Technology in Computer Science, Volume 4, Issue 3, May-June 2015.

[19]. Ionuț Taranu, “Data mining in healthcare: decision making and precision”, Database Systems Journal vol. VI, no. 4/2015.

[20]. Priyanka Sanjay Podutwar and Prof. Ms. R. R. Tuteja, “Enhancing technique for Predictive Grading of Childhood Autism using Soft Computing”, International Journal of Research In Science & Engineering, Special Issue: Techno-Xtreme 16.

[21]. Ashmeet Singh and R Sathyaraj, “A Comparison Between Classification Algorithms on Different Datasets Methodologies using Rapidminer”, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 5, Issue 5, May 2016.

[22]. D. Cutler, T. Edwards Jr., K. Beard, A. Cutler, K. Hess, J. Gibson, and J. Lawler. Random forests 6 for classification in ecology. Ecology, 88(11), 2007.

[23]. H. Drucker, D. Wu, and V. Vapnik. Support vector machines for spam categorization. IEEE Transactions on Neural Networks, 10(5), 1999.

[24]. J. R. Quinlan. Induction of Decision Trees. Machine Learning, 1(1):81{106, 1986.

[25]. X. Wu, V. Kumar, Q. J. Ross, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, P. S. Yu, Z. H. Zhou, M. Steinbach, D. J. Hand, and D. Steinberg. Top 10 algorithms in data mining, volume 14. 2008

[26]. “Rapid Miner tool”, http://rapidminer.com/

[27]. “Vaccine Adverse Event Reporting System” (VAERS), https://vaers.hhs.gov/data/data.

[28]. Priti S. Patel and Dr. S.G. Desai, “A Comparative Study on Data Mining Tools”, International Journal of Advanced Trends in Computer Science and Engineering, Vol.4(2), March - April 2.