International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 7, Issue 7, July 2017)

Prognosis of ADHD using R Programming and MATLAB Tools

M. Kartheeswari[1], K. Krishnaveni[2], E. Radhamani[3]
[1]M.Phil Scholar, [2]Associate Professor, [3]Research Scholar, Sri S.R.N.M College, Sattur, Tamil Nadu, India.

***Abstract*—Attention Deficit Hyper Activity Disorder is a Neurological Disorder characterized by problems with attention, impulsivity, and hyperactivity. ADHD may occur due to many causes .The main aim of this research work is to identify the factors that influence ADHD among the school going children. In this work initially the data samples are collected based on ADHD symptoms, brain damage and ADHD risk factors questionnaires from mind care hospitals. The data samples are pre-processed and classified into different levels of ADHD by C5.0 Decision tree classifier. The data samples classified as severe level is further classified into different types of ADHD (Inattention, Hyperactivity and Impulsivity). These data samples are checked against the attributes of brain damage and risk factor questionnaire and proved that the severe ADHD data samples having high level of brain damage and risk factor values. The experimental analysis concludes that the child will be influenced with ADHD if it has any brain abnormalities and birth risk factors. The results are analyzed by means of confusion matrices with the help of R Studio and MATLAB.**

***Keywords*—ADHD, Confusion Matrix, Hyperactivity, Impulsivity, Attention, Brain Damage, Risk Factor, Decision Tree, R Studio and MATLAB.**

I. INTRODUCTION

ADHD is a complex childhood mental disorder marked by a stable pattern of inattention and/or hyperactivity-impulsivity that interferes with functioning or development. ADHD symptoms can change over time as a person ages. ADHD hyperactivity-impulsivity is the most predominant symptom in school going children which cause the child to struggle academically. Boys are about three times more likely than girls to be diagnosed with it. Besides various causes for ADHD, brain damage or brain injury will be the main Risk factor. ADHD can make children more prone to traumatic brain injuries, and that severe enough injuries can also contribute to a form of ADHD [7].

The main aim of the work is to identify the presence and influence of ADHD due to brain damage and there by projecting the need and importance of the development of ADHD diagnostic tool to create awareness about this disorder.

In our previous work, the data samples are classified into NOADHD and ADHD with different levels by J48 and Naïve Bayes classifiers.

The experimental results show that J48 classifier gives 100% accuracy compared to Naïve Bayes which yields 94.3% with less time [1].

J.Anuradha et.al, proposed Classification Rules for Attention Deficit Hyperactive Disorder using Decision tree, BF tree, Decision Stump, AD tree and J48. The Decision tree classifier gives better accuracy with less computational time [2].

The purpose of the research work is to propose the classification algorithms to extract the ADHD data samples due to brain abnormality. This paper is organized into IV sections. The proposed methodology is discussed in Section II. The experimental results and analysis is carried out in Section III. The conclusion and the future direction of work are made in Section IV.

II. PROGNOSIS OF ADHD USING R PROGRAMMING AND MATLAB TOOLS

The Proposed research work to identify the influence of Brain Abnormality in ADHD Patients is carried out in three phases: collection of data samples, classification of data samples and prediction of disorder with brain abnormality. The data samples of school going children in the age group of 2-15 are collected from various mind care hospitals and classified initially into ADHD and NOADHD samples. Then the ADHD samples are further classified by various risk factors and brain damage symptoms. This classification will extract the ADHD samples due to brain abnormality or brain damage. The schematic diagram of this research work is shown in Figure 1.

A. Preparation of Questionnaire

The primary task of this research work is to prepare a questionnaire based on Symptoms of ADHD, Brain Damage, and Risk Factors.

1) ADHD Symptoms Questionnaire

The ADHD symptoms questionnaire proposed in [1] is used to first classify the data samples into ADHD and NOADHD levels. The principle characteristics of ADHD are inattention, hyperactivity, and impulsivity. ADHD symptom Questionnaire consists of 30 items rated on a four point scale from 0 to 3(0-Never, 1-Rare, 2-Often, 3-Very Often).

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 7, Issue 7, July 2017)

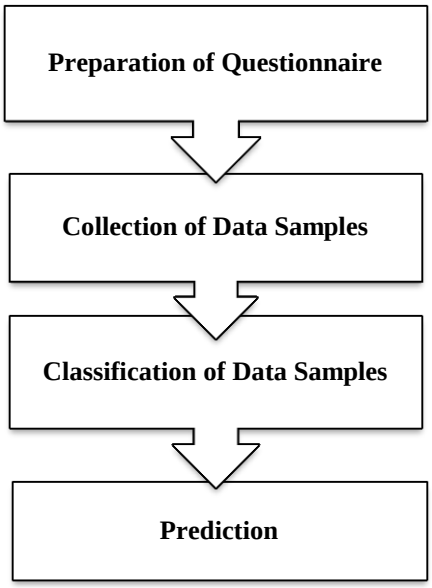


Fig 1. Schematic Diagram

2) ADHD Risk Factor Questionnaire

In neuro science, Antenatal, prenatal, perinatal and postnatal factors are considered as risk factors for the cause of ADHD. The ADHD risk factor questionnaire is prepared by considering these factors. It consists of 24 items rated on a two point scale 0 and 2(0-No, 2-Yes).

Antenatal Risk Factor

The Risk factors before birth are termed as Antenatal factors which are given in Table I.

TABLE I:
ADHD RISK FACTORS OF ANTENATAL QUESTIONNAIRE

Antenatal Risk Factors
1. Past mental health history of physical (including domestic violence), emotional abuse.
2. Current level of supports.
3. Relationship with mother and partner.
4.Anxiety
5. Any other ill treatment or tablet in pregnancy time.
6. Stressors in the last year (including bereavement, separation etc.)

Prenatal Risk Factor

The risk factors during the prenatal period (9^{th} to 10^{th} months of human pregnancy) are termed as prenatal risk factors which are given in Table II.

TABLE II:
ADHD RISK FACTORS OF PRENATAL QUESTIONNAIRE

Prenatal Risk Factors
1. Maternal Age
2. Parents are relatives
3. Family members / Siblings
4. Prenatal marital discord.
5. Lower socio-economic status / Poverty
6. Chemical exposure / Working in Pollution area / misusing drugs while pregnant

Perinatal Risk Factor

The risk after birth is identified as perinatal risk factors which are given in Table III.

TABLE III:
ADHD RISK FACTORS OF PERINATAL QUESTIONNAIRE

Perinatal Risk Factors
1. Home delivery
2. Emergency caesarean
3. Forceps delivery
4. Prematurity
5. Low birth (Birth Weight < 2.5 kg)
6. Neonatal jaundice / Sepsis / Convulsion

Postnatal Risk Factor

The Postnatal period means 1 to 6 weeks after birth. The postnatal risk factors are given in Table IV.

TABLE IV:
ADHD RISK FACTORS OF POSTNATAL QUESTIONNAIRE

Postnatal Risk Factors
1. Genetics
2. Infection exposure during pregnancy or at time of birth.
3. Brain damage either in the womb or in the first few years of life.
4. Children with head injuries after birth.
5. Food additives (dairy / artificial colouring / sugar / preservatives).
6. Nutritional deficiencies.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 7, Issue 7, July 2017)

3) Brain Damage Questionnaire

The Brain damage Questionnaire consists of 15 items rated on a four point scale from 0 to 3(0-Never, 1-Rare, 2-Often, 3-Very Often) is given in Table V.

TABLE V:
BRAIN DAMAGE QUESTIONNAIRE

Symptoms	
1. Headaches	9. Trouble with attention or thinking
2. Vomiting	10. Loss of coordination
3. Convulsions	11. Confusion
4. Abnormal dilation of the eyes	12. Aggressive
5. Inability to awaken from sleep	13. Abnormal Behaviour
6. Weakness in extremities	14. Slurred Speech
7. Trouble with memory	15. Coma or Other disorders of consciousness
8. Trouble with concentration	

B. Collection of Data Samples

The primary data samples were collected from various mind care hospitals after the preparation of questionnaires. The ADHD children were directly interviewed and an ADHD dataset with 44 samples was collected. The collected data is entered in Microsoft Excel Spreadsheet and classified by R Programming and MATLAB Research tools.

C. Classification of Data Samples

Classification is a great information mining strategy based on machine learning. To predict the target class is the main goal of classification. Initially, the data samples are classified into ADHD and NOADHD by C5.0 R programming algorithm. And ADHD data samples are further classified into different levels of ADHD (Mild, Moderate and Severe).

C5.0 Classification

C5.0 is an extension of C4.5 algorithm which gives similar results with reduced computational time. This algorithm is used to extract informative patterns, multi-value attributes and missing attributes from ADHD training data set. Then the decision tree rules are created for classification.

1) Classification: ADHD Symptoms

The ADHD symptom data set contains Name, Age, Gender, Inattention, Hyperactivity, Impulsivity, TRV (Total Rating Scale Value) and ADHD Level. The attribute ADHD Levels is determined based on TRV.

The Rules for ADHD Symptoms,
If TRV <10 ADHD Level=Mild
If TRV >=10&TRV <30 ADHD Level=Moderate
If TRV >=30 ADHD Level=Severe

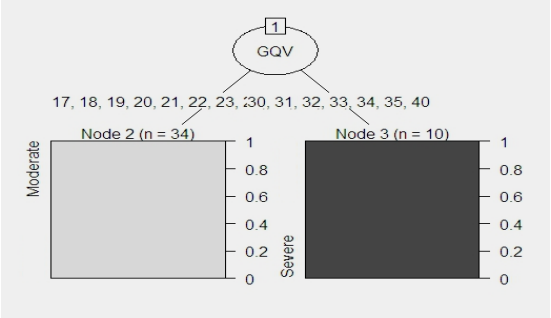


Fig 2. Decision Tree for ADHD Dataset

The decision tree used by C5.0 classifier for ADHD Symptoms Data set is shown in Figure 2. If the TRV is greater than or equal to 30 then the data is said to have Severe ADHD. Otherwise if the TRV is greater than or equal to 10 and less than 30 then the data is said to have Moderate ADHD. Otherwise, the ADHD Level is Mild.

2) Classification: ADHD Risk Factors

The ADHD risk factor data set contains Name, Age, Gender, Antenatal, Prenatal, Perinatal and Postnatal, RFV (Risk Factor Value) and RF Level. The attribute ADHD Level can determined based on RFV.

The Rules for ADHD Risk Factors,
If RFV <10 RF Level=Mild
If RFV >=10 & RFV<20 RF Level=Moderate
If RFV >=20 RF Level=Severe

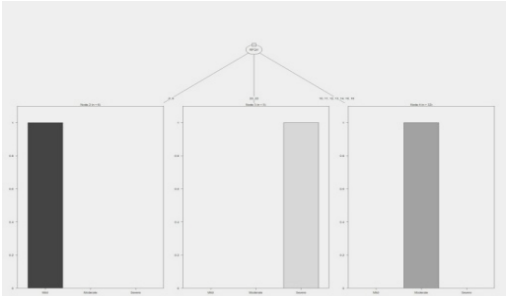


Fig 3. Decision Tree for Risk Factor Dataset

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 7, Issue 7, July 2017)

The decision tree used by C5.0 classifier for ADHD Risk Factor Data set is shown in Figure 3. If the RFV is greater than or equal to 20 then the data sample is said to have Severe ADHD. Otherwise if the RFV is greater than or equal to 10 and less than 20 then the data is said to have Moderate ADHD. Otherwise, the RF Level is Mild.

3) Classification: Brain Damage

The Brain damage dataset contains Name, Age, Gender, BDRV (Brain Damage Rating Scale) and BD Level. The attribute Brain Damage Level can determined based on BDR value.

The Rules for Brain Damage Level,
If BDRV <10 BD Level=Mild
If BDRV >=10 & BDRV <15 BD Level=Moderate
If BDRV >=15 BD Level=Severe

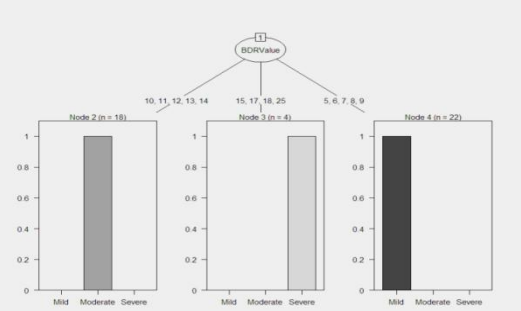


Fig 4. Decision tree for Brain Damage Dataset

The decision tree used by C5.0 classifier for Brain Damage Data set is shown in Figure 4. If the BDRV is greater than or equal to 15 then the data sample is said to have Severe BD. Otherwise if the BDRV is greater than or equal to 10 and less than 15 then the data is said to have Moderate BD. Otherwise, the BD level is Mild.

III. Experimental Analysis & Results

In this research work, R programming and MATLAB tools are used to predict the brain abnormality. It contains the following steps:

- Data Pre-processing and Attribute Selection
- Classification
- Prediction

A) Data Pre-processing and Attribute Selection

The three datasets, Namely ADHD Symptoms, Risk Factors of ADHD, Brain Damage, with 44 instances are given as input to R Studio tool.

The Sample Datasets are pre-processed by R Programming to remove duplicate records, missing and inconsistent data. After pre-processing attributes reduction is made to improve the computational efficiency. The attributes for three data sets are given in Table VI, VII and VIII.

Table VI:
ADHD Symptoms Data Set Attributes

Number	Name	Type
1	Gender	Numeric
2	Age	Numeric
3	Inattentive	Numeric
4	Hyper	Numeric
5	Impulsive	Numeric
6	TRV	Numeric
7	ADHD Level	Nominal

Table VII:
ADHD Risk Factor Data Set Attributes

Number	Name	Type
1	Gender	Numeric
2	Age	Numeric
3	Antenatal	Numeric
4	Prenatal	Numeric
5	Perinatal	Numeric
6	Postnatal	Numeric
7	RQV	Numeric
8	RF Level	Nominal

Table VIII:
ADHD Brain Damage Data Set Attributes

Number	Name	Type
1	Gender	Numeric
2	Age	Numeric
3	BDRV	Numeric
4	BD Level	Nominal

B) Classification

Classifier in R Studio is the model for predicting the nominal and numeric Quantities. Initially C5.0 algorithm is applied on the three datasets (ADHD Symptoms, Risk Factor and Brain Damage Data Set). It displays the pruned tree structure format. The data samples are classified and the classifier results are shown in Figures 5, 6 and 7.

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 7, Issue 7, July 2017)

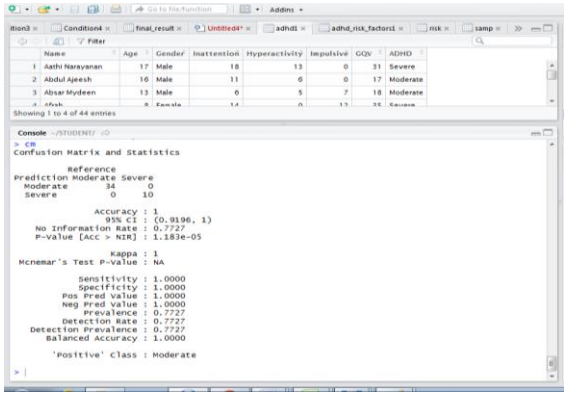


Fig 5. Output of C5.0 for ADHD Symptoms

In Figure 5, C5.0 algorithm classifies the ADHD data samples accurately. From the result it has been identified that out of 44 children 34 have moderate level of ADHD and 10 have severe level of ADHD.

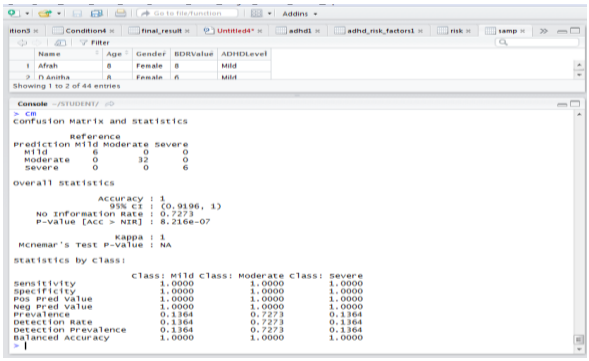


Fig 6. Output of C5.0 for ADHD Risk Factors

From Figure 6, C5.0 algorithm classifies the ADHD data samples accurately. Out of 44 children 6 have mild level risk, 32 have moderate level risk and 6 have severe level risk.

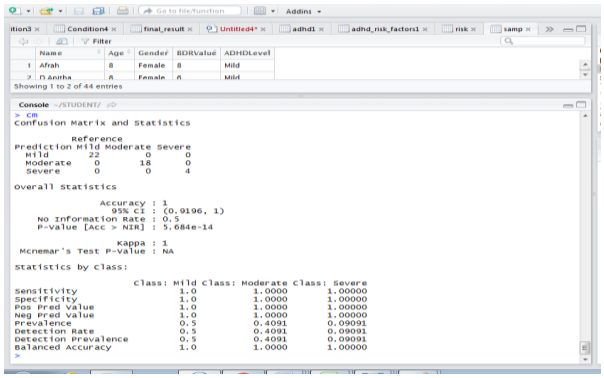


Fig 7. Output of C5.0 for Brain Damage

From Figure 7, it has been identified that 22 have mild level brain damage, 18 have moderate level brain damage and 4 have severe level brain damage out of 44 children.

In this research, the classifier results are analysed by means of confusion matrix. The high symptoms ADHD children are predicted by MATLAB Tool. And then, how far ADHD occurs due to brain abnormality is extracted by means of R Programming.

Confusion Matrix

A confusion matrix is used to describe the performance of a classification model. In classification problem, confusion matrix is a primary source of the performance measurement. The confusion matrices on three datasets results are shown in Figures 5, 6 and 7. From the result, three types of data samples (ADHD Symptoms, Risk Factor and Brain Damage) have 95% of accuracy.

C)Prediction

After the data samples are classified into various levels of ADHD, MATLAB Tool is applied to identify the significance of brain abnormality among the ADHD children i.e., the children will be predicted to have ADHD if they have brain damage or brain abnormality.

MATLAB Tool Prediction

From the results of C5.0 algorithm, the children affected with severe ADHD are taken as input, classified into three different types of ADHD by using MATLAB Tool and the results are shown in Figure 8, 9 and 10 respectively. It has been recognized that out of 44 children, 4 have severe ADHD with inattention, 5 have severe ADHD with impulsivity and 6 have severe ADHD with hyperactivity.

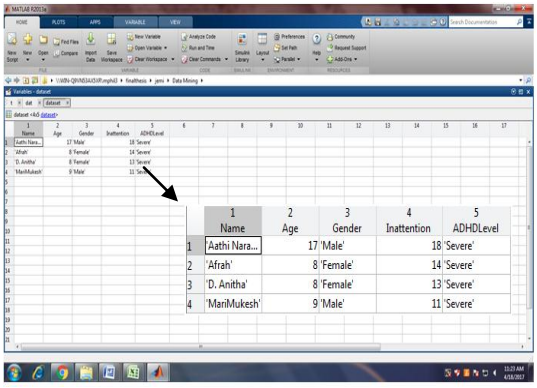
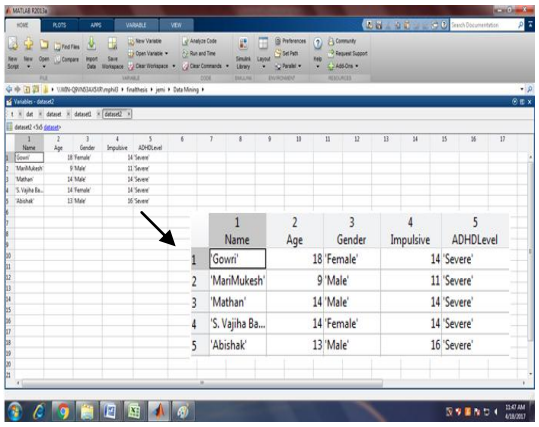


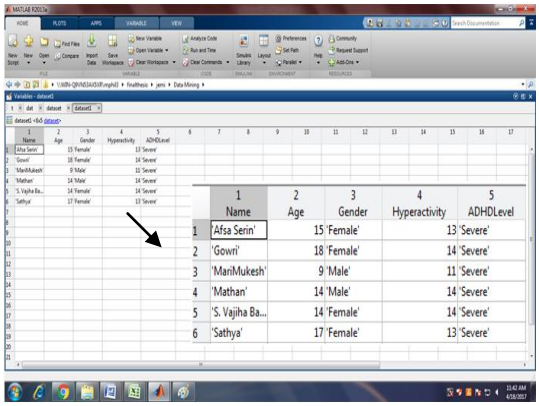
Fig 8. Severe with Inattention ADHD Children

International Journal of Emerging Technology and Advanced Engineering
Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 7, Issue 7, July 2017)



	1 Name	2 Age	3 Gender	4 Impulsive	5 ADHDLevel
1	'Gowri'	18	'Female'	14	'Severe'
2	'MariMukesh'	9	'Male'	11	'Severe'
3	'Mathan'	14	'Male'	14	'Severe'
4	'S. Vajiha Ba...	14	'Female'	14	'Severe'
5	'Abishak'	13	'Male'	16	'Severe'

Fig 9. Severe with Impulsivity ADHD Children

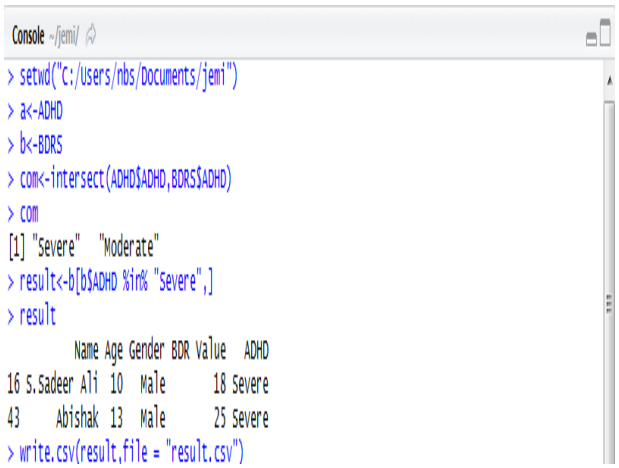


	1 Name	2 Age	3 Gender	4 Hyperactivity	5 ADHDLevel
1	'Afsa Serin'	15	'Female'	13	'Severe'
2	'Gowri'	18	'Female'	14	'Severe'
3	'MariMukesh'	9	'Male'	11	'Severe'
4	'Mathan'	14	'Male'	14	'Severe'
5	'S. Vajiha Ba...	14	'Female'	14	'Severe'
6	'Sathya'	17	'Female'	13	'Severe'

Fig 10. Severe with Hyperactivity ADHD Children

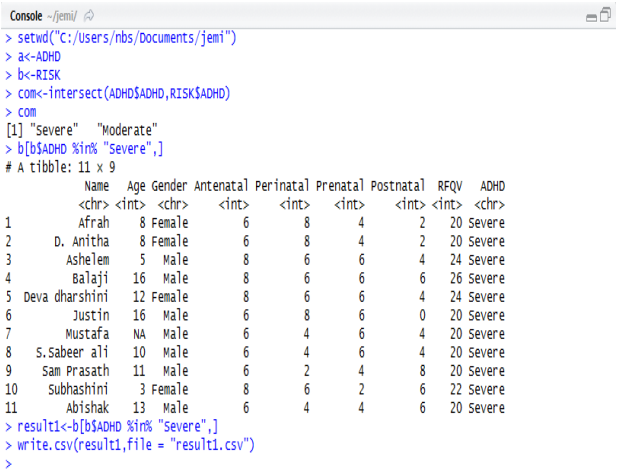
R Studio Prediction

In this phase, R Programming tool is used to find out the causes of ADHD by using Brain damage and risk factor data set attributes. The data samples labelled as severe ADHD (TRV>30) by ADHD symptoms questionnaire are checked against the brain damage attributes questionnaire. The severe ADHD samples are found to have high BDRV (Severe brain damage). From the figure 11 it has been identified that the data samples 16 and 43 having severe ADHD are also having high BDRV. The extracted severe ADHD with severe brain damage attributes are further checked against the attributes of the risk factor questionnaire. Figure 12 indicates that the data samples having risk factors before, during, and after birth can cause ADHD.



```
Console ~/jemi/
> setwd("C:/Users/nbs/Documents/jemi")
> a<-ADHD
> b<-BDRS
> com<-intersect(ADHD$ADHD,BDRS$ADHD)
> com
[1] "Severe"   "Moderate"
> result<-b[b$ADHD %in% "Severe",]
> result
         Name Age Gender BDR Value  ADHD
16 S.Sadeer Ali  10   Male       18 Severe
43      Abishak  13   Male       25 Severe
> write.csv(result,file = "result.csv")
```

Fig 11. Brain Damage ADHD Children



```
Console ~/jemi/
> setwd("C:/Users/nbs/Documents/jemi")
> a<-ADHD
> b<-RISK
> com<-intersect(ADHD$ADHD,RISK$ADHD)
> com
[1] "Severe"   "Moderate"
> b[b$ADHD %in% "Severe",]
# A tibble: 11 x 9
           Name   Age Gender Antenatal Perinatal Prenatal Postnatal  RFQV   ADHD
          <chr> <int>  <chr>     <int>     <int>    <int>     <int> <int>  <chr>
1          Afrah     8 Female         6         8        4         2    20 Severe
2      D. Anitha     8 Female         6         8        4         2    20 Severe
3        Ashelem     5   Male         8         6        6         4    24 Severe
4         Balaji    16   Male         8         6        6         6    26 Severe
5  Deva dharshini    12 Female         8         6        6         4    24 Severe
6         Justin    16   Male         6         8        6         0    20 Severe
7        Mustafa    NA   Male         6         4        6         4    20 Severe
8    S.Sabeer ali    10   Male         6         4        6         4    20 Severe
9     Sam Prasath    11   Male         6         2        4         8    20 Severe
10     Subhashini     3 Female         8         6        2         6    22 Severe
11        Abishak    13   Male         6         4        4         6    20 Severe
> result1<-b[b$ADHD %in% "Severe",]
> write.csv(result1,file = "result1.csv")
>
```

Fig 12. Risk Factors of ADHD Children

From these experiments, it has been proved that if a child has any brain damage or injury or affected with any risk factors listed in the table of I, II, III and IV respectively. It may definitely have an ADHD.

IV. Conclusion

Initially three different questionnaires based on ADHD symptoms, brain damage and ADHD risk factors are prepared and 44 ADHD data samples are collected from various mind care hospital all over Tamil Nadu.

The data samples are pre-processed by R Programming techniques and C5.0 Decision tree classifier is proposed to classify the children into mild, moderate and severe level of ADHD. Once the ADHD samples are classified, MATLAB Tool is applied to further classify the severe ADHD samples into its different types. Then R Programming is used to find out the causes of ADHD with the help of brain damage and risk factor questionnaire. The results show that the severe ADHD data samples are also having high brain damage and risk factors. So it has been concluded that if a child has any brain abnormality and the risk factors taken in this research work will definitely have ADHD. Since medication and behavioural therapy requires lot of attention, the research may be directed towards the implementation of diagnosis tool by using brain imaging techniques to redemption from this disorder.

REFERENCES

[1] K.Krishnaveni and E.Radhamani, "Diagnosis and Evaluation of ADHD using Naïve Bayes and J48 Classifiers", 2016 International Conference on Computing for Sustainable Global Development (INDIA Com), 978-9-3805-4421-2/16/$31.00_c 2016 IEEE.

[2] J.Anuradha,N.Jayasri,K.V.Arulalan, "Classification Rules for Attention Deficit Hyperactive Disorder using Decision Tree", International Journal of Computer Science and Communication Networks,Volume2(2), 295-303, ISSN:2249-5789.

[3] K.Krishnaveni and E.Radhamani, "Diagnosis and Evaluation of ADHD using MLP and SVM Classifiers", Indian Journals of Science and Technology, Volume 9(19), DOI: 10.17485/ijst/2016/v9i19/93853, May 2016 ISSN (Online): 0974-5645.

[4] Rahman, Rashedur M., and FarhanaAfroz. "Comparison of various classification techniques using different data mining tools for diabetes diagnosis." Journal of Software Engineering and Applications 6.03 (2013): 85.

[5] R. Revathy, R. Lawrence "Comparative analysis of C4.5 and C5.0 Algorithms on Crop Pest Data", International Journal of Innovative Research in Computer and Communication Engineering, Vol. 5, Special issue 1, March 2017, ISSN: 2320-9801.

[6] Christopher M. Bonfield,M.D., Sandi Lam,M.D., Yimo Lin,B.A., "The impact of attention deficit hyperactivity disorder on recovery from mild traumatic brain injury", J Neurosurg: Pediatrics/Volume 12/August 2013.

[7] http://healthland.time.com/concussions-may-cause-more-brain-damage-in-kids-with-adhd/

[8] https://cran.r-project.org/doc/contrib/usingR.pdf.

[9] https://en.wikipedia.org/wiki/R_(programming_language).