

International Journal of INTELLIGENT SYSTEMS AND APPLICATIONS IN

ENGINEERING

ISSN:2147-6799

www.ijisae.org

**Original Research Paper** 

# Determination of Diagnostic Factors for Different Domains Using FDMML Techniques

# R. Suguna<sup>1</sup>, R. Uma Rani<sup>2</sup>

Submitted: 20/02/2023	<b>Revised:</b> 16/04/2023	Accepted: 08/05/2023
		<b>.</b>

**Abstract:** There are numerous factors that contribute to each problem in research criteria. Diagnose problems using diagnostic analytics to find causes and solutions. DA supports finding solutions and advancing science for the benefit of future generations. In this paper, there are three kinds of data have selected, which are essential and solutions needed in the future. Autism is one of the medical data which is increasing every year in India. Rainfall data is an agricultural data, which is a main need for earth. Child abusing is the social relevant data, which should be considered earnestly in our India. Diagnostic analytics can be attained through some of statistical techniques such as correlation, hypothesis testing and regression techniques. The three types of data's proposed factors are analysed using historical research and computational intelligence methodologies. Effective fuzzy rules are generated for the diagnosed factors.

Keywords: Diagnostics analytics, correlations, machine learning, autism, rainfall, child abuse

# 1. Introduction

#### 1.1 Diagnostic analytics in data science

Data science is a bundle of data analytics techniques, machine learning techniques and mathematical techniques. Analytics plays a significant part in the broad field of data science to address a variety of challenges with the collaboration of machine learning techniques. Regression and correlation can be providing fruitful results in the zones of diagnostic analytics and prediction analytics. Different types of analytical processes that can be applied to data are provided by data analytics. The sub parts of data analytics are such as descriptive, diagnostic, predictive and prescriptive analytics. Diagnosing a problem is the way of investigating reasons, which is answer for why it happened? .Diagnostic analytics is also called as root cause analysis. It is deeper dump into data than the descriptive analytics. Techniques for diagnostic analytics such as hypothesis testing, correlation, causation and diagnostic regression can be used for analysing reasons and connections between the attributes (Catherine Cote, 2021).

#### **1.2 Machine learning techniques**

Machine learning techniques are the efficient and fast due to the comparison of other methods. Simple techniques like frequency of distribution, correlation, standard deviation and summary of data, covariance are be part of descriptive

1Research Scholar, Sri Sarada College for Women (Autonomous), Salem-16. sugunarmca@gmail.com 2Principal Sri Sarada College for Women (Autonomous), Salem-16. principalsscsalem@gmail.com analytics. Predictive data analytics make use of regression and time series methods for better outcomes .Causation is statistical methodology, which derives association effectively. Diagnostic analytics also used correlation and hypothesis testing to demonstrate the connection between the variables. Prescriptive analytics is using optimization techniques to get efficient results from data.

#### 1.3 Fuzzy System

Fuzzy system is a section of computational intelligence environment. It is a very adaptable area for efficiently handling various forms of data. Fuzzy systems are widely used in all the domains like engineering, biological problems, financial and other areas also. Decision making is the need to improving the current scenario and identifying faults in every era. A fuzzy rule delivers more worthy ideas and simple to understand than the other methods. (Dragan Pamucar, 2022)

#### 2. Literature Review

#### 2.1 Autism

Autism is a frightening brain nerve problem in the world wide. It is affecting the brain neuron at any stage of human based on variety of reasons. This disorder disturbs the communication, sleeping, repetition of behaviours and making depression in human. Mostly autism affects children in the infant stage. But it is not identified easily before the age of 2. Autism is affecting the male, comparing with the affecting rate of female. April 2 is observed that the world's autistic awareness day (Laura A. McGuinn, 2020). Top countries with highest rate of autism among the world is, Table.1: Ranking of countries based on autism rate

Position	Country name	Rate(per 10,000)
1	Qatar	151
2	United Arab Emirates	112
3	Oman	107
4	Bahrain	103
5	Saudi Arabia	100
6	Kuwait	97
7	Jordan	92
8	Syria	91
9	Afghanistan	91
10	Palestine	91

(Source: world population review)

Many of the authors concentrated on the factors of autism and described in different ways. Genetic, epigenetic and unusual food habits were identified, then level of autism in India, symptoms of autism also discussed (sahermehdi, 2022) (Rylaarsdam, 2019).Socio-demographic and lowincome populations are recognised as aspects of autism. Research findings were collected from 2012 to 2021 along with the keywords autism and prevalence and 132 articles were considered for the review (Jinan Zeidan, 2022). Jaundice was taken as an aspect which affects children and creates a chance for ASD by their birth. Based on metaanalysis, there are 32 studies which are published articles of Scopus, PubMed, Embase, Google scholar and Cochrane till 2019.Systematic review made with the PRISMA guidelines and found that high risk bios is convicted that jaundice is associated with ASD (Monica L. Kujabi ,2021).

Based on the report of etiology, about 8000 children were affected in Jordan. The etiology report conveyed that genetic is a main factor for autism. Because damaged gene can affects infant's brain neuron (Emad, 2019). Study conducted in polluted areas of United States and 674 children were pretentious among 855, for 2013 to 2016. The study conveyed that air pollution as an issue for early life exposure and ASD. Positive association of factor was calculated with 95% of CI (Laura A, 2019).Study steered on 3964 children about 6-9 years among five regions of India. Cross sectional survey has been applied on the population of cities in India. Based on the study, child nutrition is one of the risk factor for autism (Arora NK, 2018). (Amir, 2015) has presented in a study that lack of physical activity can be a factor for autism. The research moved on 83 children with the age covered between 6-15 years. (Helen, 2011) were explained that environmental

toxicants, co-morbidities and infections are the factors of ASD.

# 2.2 Rainfall

Rainfall is International level need for people. Southwest monsoon is the main source of the rainfall in India. Indian summer monsoon rainfall (ISMR) getting 80% of the precipitation at the time of June to September (Subimal et.al, 2016). Based on the report of times of India 2021, only 236 districts of India have showered above normal rain out of 703 districts. IMD (Indian meteorological department) maintaining the reports of Indian rainfall and long time period average (LPA) of rainfall. Based on the LPA, IMD suggested that 89cm is standard rainfall. Reason for the rainfall is climate change, dry and wet epochs in India defined by IMD. India is in the 3<sup>rd</sup> place based on the forest cover among the world -wide (Times of India, 2022).

Indian regions such as peninsular, central India, Himalayan regions and northeast regions were focused and analysed for rainfall discrepancy. Each region had above normal rainfall in different monsoons depending on the topographical structures. Global warming and climate change in different regions are the reasons for rainfall variability (Hamza, 2022). Coupled Model inter comparison Project Phase (CMPI6) is a project regarding rainfall of india.CMPI6 compared rainfall and global mean temperature. Result of the work explored that global warming is the main factor for variation in rainfall of India (Anja Katzenberger, 2021).

Statistical analysis done for Indian rainfall data from1901-2020 and it is divided as quad-decades .Analysis explained as climate variability is the factor for changeability in rainfall (Abhilash, 2021). Another findings were analysed the decadal rainfall for years (1951-1984) and (1985-2018).Continuous rainfall for 10 days and 11 days in the decadal years compared and analysed that the significant level of trend analysis for the Indian summer monsoon is not smooth. Man-kendall test is applied for the trend analysis. The study revealed that the factor for abnormality of summer monsoon in India is air pollution, diminishing monsoonal circulation and heating of the Indian Ocean (Vinay Kumar, 2021). Report from ministry of earth sciences showed about the decade (1986-2015) and past history of seasonal, monthly and annual rainfall description based on statistical techniques. Main factors for infrequent rainfall such as global temperature rise, ocean warming, droughts, sea level rise and tropical cyclones and changes in Himalayas (R.krishnan, 2021). Forecasting of rainfall was performed with the help of ARIMA, ARMA, and SARIMA. Analysis of annual mean rainfall and change point was calculated with the pettit test for the period (1979-2015). Low convective rainfall rate makes high convective precipitation .So; it provides more than normal

moisture in Indian Ocean, which is an unfavourable condition for abnormal rainfall (Bushra, 2020). The daily gridded rainfall for India has evaluated using mann-kendall test to analyse trends of rainfall. Data has taken for the period (1901-2004). This study explored the causes for abnormal rainfall is climate change. By this factor wet areas becomes moistened and dry areas becomes patched (Subimal, 2016).

### 2.3 Child abuse data

World Health organization refers that child abusing means different kinds of problems facing between the ages of 0 to 18 years. Studies show that, every 3 in 4 members were affected by the psychological fierceness on the age of 2 to 4 years of children. Every year 40-150 deaths occurred worldwide, under the age of 18 years due to the child maltreatment. Most of the children were facing variety of problems (WHO, 2022). (Vinoth paul, 2021) described that abusing is an evil of society. Care takers, doctor, parents should create awareness regarding such painful situations to children.

Child protection team should be allotted to the areas, which is having high density of problems. Mental illness, parental obstruction, disharmony are the factors which are conceded by the author. Study was concentrated on awareness of child sexual abuse (CSA). Results revealed that statistically most of people are aware about POCSO (Protection of children from sexual offences) comparing with online compliant system of child abuse and one-stop centre at hospitals. Lack of awareness and lack of education is needed in the society (Ranjana, 2021). Frequency of child abuse has reviewed and electronic bibliographic data bases such as Cumulative Index to Nursing and Allied Health (CINAHL), MEDLINE, global health, Scopus and academic search complete. Reference papers also scanned for the study in the period of June to September 2021.there are 150 paper regarding child abuse were collected and 20 papers selected based on the tools of Effective Public Health Practice Project (EPHPP) and Critical Appraisal Skills Programme (CASP) .The study

explored that up to 150 deaths occurred every year due to child abuse. Boys and girls in the range of 17% and 31% of were suffered by different kinds of child abusing problems. The study discussed that high frequency of child abusing is occurring in slum areas and impression rooted in poverty and socioeconomic situations of the children (Nishamary, 2021).

Report of IAP briefly explained about child abusing. Types of abuses, abusing methods, reasons for abuser, POSCO act, Illiterate and addiction for drugs are the main causes of child abuse (Kiran, 2021). (Amisha, 2019) showed that the, types of child abusing, causes and awareness among the society regarding child abuse were explained. The Study revealed that addiction to alcohol, drugs and personnel tribulations are the sources for child abusing. Abusing and neglecting of 720 adolescents in the place Jammu was analysed using MIMIC model. Illness of parental education is a big factor on abusing and neglect of adolescents based on the correlation (Rubycharak, 2014)

# 3. Methodology

The overall methodology of this paper is described in the following picture. In this paper, diagnosing factors for different domains are specified using two various measures such as,

- Based on the studies of the domains
- Based on machine learning techniques



Fig.1. Diagnosing factors for different domains

# 3.1 Correlation

Correlation is a technique which is describing the association among two variables. There are positive and negative correlations are available based on the data. Based on the correlation, autism and bilirubin data is compared and in the same way annual rainfall and tree density will be compared for the amount of association among the attributes.

# **3.2** Causation

Causation is methodology for expressing association among the variables. Causation is not like the correlation. When two attributes are taking for the causation, effect of one attribute can affect another attribute. This is known as causation connection of the variables. Significance level which is an estimator in the hypothesis testing can be justified the causation among the variables. The relation of causation of factors for autism, rainfall and child abuse data is defined as,

- ▶ IF Jaundice is high , then autism is low
- ▶ IF Tree density is high then rainfall is high
- ➢ IF AD is High then child crime is high

# 3.3 Chi-Square test

Chi-Square test is an effective way for comparing two variables based on some hypothesis. Especially categorical data can be analysed using chi square test. In this paper, autism data is having categorical data. So, chi-square test is appropriate technique for autism data.

# **3.4 Diagnostic regression analysis**

Regression analysis helps for diagnostic analytics to take a decision on the bonding among the variables. Linear regression is applied on the three types of data in the following manner.

• Bilirubin level of jaundice and result of autism

- Tree density of states of India and annual rainfall
- Drug addicts in states of India and state wise child abuse statistics

#### 3.5 Identified factors based on the studies:

#### Autism: Jaundice

- Autism-jaundice: Jaundice is a harmful problem in childhood. Bilirubin is a yellow colour pigment in the blood bile and which is main cause of jaundice. Range of direct bilirubin is 0.3ml/dl and total bilirubin is 0.1 to 1.2 mg/dl in the blood. When bilirubin gets increasing, then jaundice can be identified this is affecting the brain easily (Yvonne, 2016).
- Based on the studies, chemical mixed food items are the factors for autism is in children. Packed food items having some harmful chemicals such as bisphenyl, perchlorate and some other chemicals which are harmful to brain system of the children (TakeoFujiwara, 2016).

# Rainfall: Trees density

Trees density for a region is describing a factor for rainfall .Because trees makes climate as smooth and reducing soil erosion. Computationally, trees cover in medium dense data is compared with annual rainfall of India for years 2001-2021.regression and correlation is applied on this data.

# Child abuse: Addiction of drugs

Based on the discussion of literature studies, most of the studies showed that the drug addiction and illeducation and drug addiction are the main causes for child abuse.

# 3.6 Proposed Diagnose factors using Computational intelligence technique:

Table.5: Procedure of Diagnostic Fuzzy Rules

Diagnostic Fuzzy Rules

Input: Statistical report of data Output: Fuzzy rules

Step 1 : Load the data D

Step 2 : Initialize the linguistic variables for factors(a1,a2..an)

Step 3 : Combine the factors and linguistic variables

Step 4 : Framing the rules using linguistic variables

#### IF a1 is LV1 && a2 is LV2

THEN Decision variable is LV

Step 5 : Display fuzzy rule

- In the above table, proposed approach is designed for different kinds of data using fuzzy rules. In the Crisp data can create only two linguistic variables such as YES and NO.
- Factors are the attributes which are related to decision making. Linguistic variables are flexible labels such LOW, MEDIUM, HIGH.
- Decision making is taking an important position in diagnostic analytics. Fuzzy rules are helpful to design the factors for the data.

#### 4. Results and discussion

#### 4.1Diagnostic regression analysis:

Data	Degrees of freedom(DF)	F-Statistic	P-value	Adjusted R <sup>2</sup> Value
Autism data	28	2.649	0.420	74.97
Rainfall data	10	7.095	0.258	78.87
Child abuse data	29	2.197	0.401	67.23

Table.6: Metrics of Diagnostic Regression Analysis

Performances of linear regression with the different kinds of data are explored in the above table.5.F-statistic is a component of evaluating the entire regression model and it should be >1. P value for all the data is >0, hence the depiction of model is suitable for the proposed factors under the 95% of confidence level. Adjusted  $R^2$  is an estimator, which pretends the variation in the data for the model.



Fig.2. Association among the diagnostic factors with autism ,rainfall and child abuse data

Fig.2 represents plot formation of linear relationship among the attributes of different kinds of data.result of autism and bilirubin range in the autism data.Tree density and annual rainfall of india are compared in the rainfall data.In child abuse data state wise child crime and statewise drug addicts are compared.



**Fig.3.** Residual q-q and leverage plots for autism, rainfall and child abuse data

In the above fig.3, residuals-Q plots, leverage and cook's distance are explored for the three kinds of data. A residual plot shows the linearity of the measures in the data. Line on the residual plot represents the quality the linearity of data. Q-Q plot discovers that the normality of data. 45° line shows that the good normality available in the diagnosing factors. Scale-location intimates that square looted of observed predicted values and it specifies the similarity of the residuals in the model. Leverage or cook's distance is a plot which shows the standard residuals of the model.

# 4.2 Chi-Square test $(\chi^2)$ and correlation





Fig.4.  $\chi^2$  of autism data and correlation plots of rainfall and child abuse data

In the above fig.4 delivers the  $\chi^2$  test on autism data, which is having categorical data. Based on the  $\chi^2$  test, jaundice is a factor for autism. Correlation plot specifies that the relationship among the diagnosed factors and different data are positively correlated.

#### 4.3 Proposed diagnostic fuzzy rules

A fuzzy system consisting of 3 variables and 9 rules.	Variables:
Variables: jaundice(L, M, H) autism(L, M, H)	rainfall(L, M, H) trees(LD, M, HD) WS(L, M, H)
chemical(L, M, H)	Rules:
Rules: jaundice %isi% H && chemical %isi% L $>$ autism %isi% M jaundice %isi% H && chemical %isi% M $>$ autism %isi% M jaundice %isi% H && chemical %isi% M $>$ autism %isi% H jaundice %isi% L && chemical %isi% H $>$ autism %isi% M jaundice %isi% L && chemical %isi% L $>$ autism %isi% M jaundice %isi% L && chemical %isi% L $>$ autism %isi% M jaundice %isi% A && chemical %isi% L $>$ autism %isi% M	$\begin{array}{l} \label{eq:constraint} trees \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ $



Fig. 5. Fuzzy rule for autism, rainfall and child abuse data

In the above fig.5 shows that the model of fuzzy rule. Autism and rainfall data are having the liguistic variables as low, medium, high. Child abuse data is having very low, medium, high, very high. Below fig.6 specifies the plots representation of fuzzy rules.



Fig.6/Plots for fuzzy rule linguistic variables

# 5. Conclusion

In this paper, diagnostic analytics performed on different kinds of data. Machine learning approaches were used to identify the diagnostic factors. Proposed fuzzy rules produced utilising tools from computational intelligence. The proposed fuzzy rules can be used to create fuzzy inference systems in the future, and different types of data can be processed utilizing perspective analytics.

#### References

- [1] Amir Hossein Memari et.al, Children with Autism Spectrum Disorder and Patterns of Participation in Daily Physical and Play Activities, Hindawi Publishing Corporation Neurology Research International Volume 2015, Article ID 531906, <u>http://dx.doi.org/10.1155/2015/531906</u>.
- [2] <u>Catherine Cote</u>, a beginner's guide to data & analytics", Harvard public school, https://online.hbs.edu/blog/post/diagnostic-analytics, 2021.
- [3] Abhilash Singh Chauhan et al, Rainfall dynamics observed over India during last century (1901–2020) using innovative trend methodology, Water Supply Vol 22 No 8, 6909 doi: 0.2166/ws.2022.291,2022.
- [4] Anja Katzenberger et al,,Robust increase of Indian monsoon rainfall and its variability under future warming in CMIP6 models, Earth System Dynamics,volume 12, https://doi.org/10.5194/esd-12-367-2021

- [5] Laura A. McGuinn et.al, Early Life Exposure to Air Pollution and Autism Spectrum Disorder Findings from a Multisite Case–Control Study, Wolters Kluwer Health, Epidemiology, Volume 31, Number 1,2020.
- [6] Yvonne W. Wu, Michael W. Kuzniewicz, MPH, Lisa Croen, "Risk of Autism Associated With Hyperbilirubinemia and Phototherapy", PEDIATRICS Volume 138, number 4 ,DOI: 10.1542/peds.2016-1813,2016
- [7] TakeoFujiwara , NahoMorisaki , YukikoHonda, MakikoSampei1 and YukakoTani , chemicals,Nutrition,andAutismSpectrumDisorder:A Mini-Review doi: 10.3389/fnins.2016.00174,2016
- [8] R. Krishnan et.al. A report of minidtry of earthsciences ,Assessment of climate change over the indian region, ,ISBN 978-981-15-4326-5 ,https://doi.org/10.1007/978-981-15-4327-2,springerOpen.
- [9] Ghosh S, Vittal H, Sharma T, Karmakar S,Kasiviswanathan KS, Dhanesh Y, et al. Indian Summer Monsoon Rainfall: Implications of Contrasting Trends in the Spatial Variability of Means and Extremes. PLoS ONE 11(7): e0158670.doi:10.1371/journal.pone.0158670,2016.
- [10] Bushra Praveen, Swapan Talukdar, Analyzing trend and forecasting of rainfall changes in India using nonparametrical and machine learning approaches ,nature research, https://doi.org/10.1038/s41598-020-67228-7,2020.
- [11] Shipra Jain, Adam A Scaife et al., Current chance of unprecedented monsoon rainfall over India using dynamical ensemble simulations, https://doi.org/10.1088/1748-9326/ab7b98, IOP publishing, 2020.
- [12] Monica L. Kujabi et al., Neonatal jaundice and autism spectrum disorder: a systematic review and meta-analysis, International Pediatric Research Foundation, Inc 2020, springer nature.
- [13] Hamza Varikoden et al., "Features of regional Indian monsoon rainfall extremes", <u>Climate Impacts on</u> <u>Extreme Weather</u>, https://doi.org/10.1016/B978-0-323-88456-3.00014-9,2020.
- [14] Medicine news today, What causes autism? Genetic and environmental factors, Dannell Roberts, PhD, BCBA-D — By Alicia Sparks Akers on October 4, 2022.
- [15] Emilio A.L.Gianicolo, Dr.physiol., et.al. Methods for Evaluating Causality in Observational Studies, Dtsch

Arztebl Int., 117(7): 101– 107, doi: 10.3238/arztebl.2020.0101,2020.

- [16] Vinod Paul et al., "Child abuse: A social evil in Indian perspective", Journal of Family Medicine and Primary Care, DOI: 10.4103/jfmpc\_jfmpc\_1862\_20,2021.
- [17] Singh R, Koushal V, Bharti B. A descriptive study on child sexual abuse act in India. Journal of Family Medicine and Primary Care, 2022; 11:2923-32.
- [18] Ruby Charak, Hans M. Koot, Abuse and neglect in adolescents of Jammu, India: The role of gender, family structure, and parental education, http://dx.doi.org/10.1016/j.janxdis.2014.06.006, Elsevier Ltd.,2014.
- [19] Kiran aggarwal,RS sethi,Guidelines for parents,when and how do I suspect and prevent abuse in child?,"Indian academy of pediatrics"IAP action plan 2021-2022.
- [20] Dragan Pamučar et al., computational Intelligence and Fuzzy Modeling for Sustainable Decision-Making in Engineering Processes,volume 2022,special issue, Hindawi publications.
- [21] Nisha Mary Rajan1, Dr. Jogindra Vati2, Prevalence of Child Abuse: A Review Study, International Journal of Science and Research (IJSR), ISSN: 2319-7064, DOI: 10.21275/SR211020074520, 2021.
- [22] Kriti Aeron, Child Sexual Abuse in India: An Understanding, International Journal of Law Management & Humanities, Volume 2, Issue 1, ISSN: 2581-5369, 2018.
- [23] Arushi panwar, Child abuse in India, International Journal of Research and Analytical Reviews, e ISSN 2348-1269, Volume 6, Issue2, 2019.
- [24] https://worldpopulationreview.com/countryrankings/autism-rates-by-country
- [25] PIB Delhi, https://pib.gov.in/PressReleasePage.aspx? PRID= 1842697,2022.
- [26] Amisha U. Pathak, Intern- India Think Council, www.indiathink.org.
- [27] https://towardsdatascience.com/from-correlation-tocausation-49f566eea954/
- [28] https://www.health.harvard.edu/blog/common-foodadditives-and-chemicals-harmful-to-children-2018072414326
- [29] https://www.who.int/news-room/factsheets/detail/child-maltreatment