DOI: 10.21917/ijsc.2017.0201

FUZZY CLUSTERING BASED BAYESIAN FRAMEWORK TO PREDICT MENTAL HEALTH PROBLEMS AMONG CHILDREN

M.R. Sumathi[1] and B. Poorna[2]
[1]*Department of Computer Science, Bharathiar University, India*
[2]*Shri Shankarlal Sundarbai Shasun Jain College for Women, India*

Abstract

According to World Health Organization, 10-20% of children and adolescents all over the world are experiencing mental disorders. Correct diagnosis of mental disorders at an early stage improves the quality of life of children and avoids complicated problems. Various expert systems using artificial intelligence techniques have been developed for diagnosing mental disorders like Schizophrenia, Depression, Dementia, etc. This study focuses on predicting basic mental health problems of children, like Attention problem, Anxiety problem, Developmental delay, Attention Deficit Hyperactivity Disorder (ADHD), Pervasive Developmental Disorder(PDD), etc. using the machine learning techniques, Bayesian Networks and Fuzzy clustering.

The focus of the article is on learning the Bayesian network structure using a novel Fuzzy Clustering Based Bayesian network structure learning framework. The performance of the proposed framework was compared with the other existing algorithms and the experimental results have shown that the proposed framework performs better than the earlier algorithms.

Keywords:

Fuzzy Clustering, Bayesian Network, Structure Learning, Prediction

1. INTRODUCTION

Mental Health Problems (MHP) have become one of the main causes of the burden of disease worldwide and it often starts at an early age and may destroy lives, with an impact on families, colleagues and communities. According to survey conducted by World Health Organization, depressive disorders will become the second leading cause of the global disease burden by 2020 [1].

There are strong links between physical and mental health problems and if untreated, these conditions severely influence children's development, their educational attainments and their potential to live fulfilling and productive lives. The best way to deal with a crisis is to prevent it from happening in the first place. Hence, in this study a new framework has been developed to help the psychologists in predicting the MHP at an early stage. Expert system (ES) is a branch of Artificial Intelligence that works within a particular domain. These are designed to provide expert quality performance on domain specific problems.

Mental Health Diagnostic Expert Systems mimic the role of a psychologist with a view to diagnose the MHP. Various Machine Learning techniques are available to develop expert systems. Bayesian Networks is a popular machine learning technique which is widely used in expert systems for predicting diseases. This machine learning technique is based on probabilities and Bayes Theorem.

1.1 BAYESIAN NETWORKS

Bayesian networks (BN), a graphical model representing probabilistic relationship among random variables, have been successfully applied in many domains like medical diagnosis, hardware troubleshooting, etc. The BN model is well suited for representing uncertainty and causality. A BN consists of two components, a network structure in the form of a Directed Acyclic Graph (DAG) and a set of Conditional Probability Distributions (CPD), one for each variable characterising the dependencies between the variables. A simple Bayesian Network is shown in Fig.1.

If there is a directed edge in a DAG from node *Y* to node *Z*, *Y* is said to be a parent of *Z* and *Z* is called a child of *Y*. The presence of an edge connecting two nodes implies that these two variables are dependent. The absence of an edge between any two nodes implies that these two variables are independent given the values of any intermediate nodes. BNs have been used in other studies for predicting various mental disorders like schizophrenia, depression, psychosis, etc. This study concentrates on predicting the basic mental health problems that affect children, viz. Attention problem, Academic Problem, Anxiety Problem, Attention Deficit Hyperactivity Disorder and Pervasive Developmental Disorder.

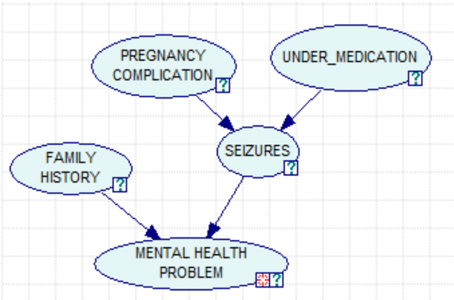


Fig.1. Simple Bayesian Network

There are 5 nodes and 4 arcs in Fig.1. Pregnancy_ Complication and Under_Medication nodes are parents of Seizures node. The arcs show that there is a direct relationship between Pregnancy_Complication and Seizures, while there is an indirect relationship between Pregnancy_complication and Mental_Health_Problem.

Learning a BN refers to data based inference of either the conditional probability parameters for a given structure or the underlying graphical structure itself. We focus here on the process of structural learning, to discover the relations between various variables and predict the mental health problem.

BN structures are learnt using score-based and constraint-based methods. The score-based methods along with a search strategy find a structure with optimal score. More importance is given for the search strategy and the scoring function. Various scores viz., Log-likelihood, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Bayesian Dirichlet Score (BDE), etc. are available for Bayesian Networks. The constraint-based methods infer structures through conditional independency tests. The constraint-based methods are generally faster than the score-based methods and give a trustworthy result when sufficient data are provided. In this paper, we propose a framework that considers the advantages of both the methods.

1.2 FUZZY CLUSTERING

Clustering involves assigning data points to clusters such that the items in the same cluster are as similar as possible, while items belonging to different clusters are as dissimilar as possible. Broadly, clustering can be divided into two namely, Non-Fuzzy clustering and Fuzzy clustering. Fuzzy clustering is a form of clustering in which each data point can belong to more than one cluster.

Fuzzy c-means algorithm, proposed by Dunn (1973) and later improved by Bezdek (1981), is an extension of K-means where each data point can be a member of multiple clusters with a membership value. Fuzzy clustering algorithm attempts to partition a finite collection of 'n' elements of $X = \{x_1,x_2,...,x_n\}$ into a collection of 'c' fuzzy clusters with respect to some given criterion. The algorithm returns a partition matrix $W = w_{i,j}$ belongs to [0,1], $i = 1,2,….,n$, and $j = 1,2,….c$, where each element $w_{i,j}$, tells the degree to which element x_i, belongs to cluster c_j.

In this paper, an attempt has been made to develop a Fuzzy-Clustering based Bayesian Network model that predicts basic MHPs which affect children. Fuzzy clustering has been performed on the variables based on the mutual information among them. The main focus of the research is to construct the model and compare its performance with other models in predicting MHP.

This framework reduces the search space as a few variables are involved in each cluster. The proposed algorithm was experimented to predict the basic MHP among children and compare the results with other algorithms. The empirical results show that the fuzzy clustering algorithm achieves good performance in prediction.

The paper has been organized as follows. Section 2 details on the related work in predicting MHPs using BNs. Section 3 discusses on the dataset and the methodology of the study. The proposed framework has been detailed in section 4. Section 5 presents a brief on the results and evaluation. Section 6 concludes and suggests on future work that can be done on the study.

2. RELATED WORK

The research on applying machine learning techniques in mental health diagnosis has started in nineteen eighties. DTREE [2] is an expert system that diagnoses DSM-IV Axis I disorders using Decision Tree techniques. INTERNIST/AUTOSCID [3] is a computerized Structured Clinical interview for DSM-IV Axis II personality disorders. Yap and Clarke [4] developed an expert system called MILP (Monash Interview for Liaison Psychiatry) using constraint-based reasoning for systematic diagnoses of mental disorders based on DSM-III-R, DSM-IV and ICD-10. In [5], Kipli, Kuryati, Abbas Z. Kouzani, and Matthew Joordens investigated performances of four Feature Selection algorithms, namely, OneR, SVM, Information Gain (IG) and ReliefF in detecting depression from structural MRI scans.

Masri and Jani in [7] used three artificial intelligence techniques viz., Rule-Based Reasoning, Fuzzy Logic and Fuzzy-Genetic Algorithm for diagnosis and suggestion of treatment plans. David [8] analyzed how artificial intelligence has been utilized in psychological practice. The current trend and future applications and their implications have been discussed.

Razzouk et al. [9] developed a decision support system for the diagnosis of schizophrenia disorders with an accuracy of 66-82%. Chattopadhyay et al. [10] developed a neuro-fuzzy approach to grade adult depression. Nunes et al. [11] introduced a hybrid method to handle diagnosis of Schizophrenia. They have integrated structured methodologies in decision support and structured representation of knowledge into production rules and probabilities.

Basavappa et al. [12] used depth first search method with backward search strategy to diagnose depression or dementia. They developed an expert system using the patient's behavioral, cognitive, emotional symptoms and results of neuropsychological assessments. Rahman et al. [13] compared various classification techniques viz. Bayesian Network, Multilayer Perceptron, Decision Trees, Single Conjunctive Rule Learning and Fuzzy Inference Systems and Neuro-Fuzzy Inference System, using different data mining tools like WEKA, TANAGRA and MATLAB for diagnosing diabetes.

Gil and Johnson proposed a system based on Artificial Neural Networks (ANN) and Support Vector Machines (SVM) that diagnoses Parkinson Disease [14]. The system has shown an increase in accuracy and a decrease in cost. Jerzy et al. tried to find efficient methods for classification of patients with mental disorder [15]. They identified that Attribute Extension Approach improves classification accuracy in the case of discretized data. Seixas and Zadrozny [16] proposed a Bayesian Network (BN) Decision Model for diagnosis of dementia, Alzheimer's disease and Mild Cognitive Impairment. Khemphila and Boonjing [17] used Multi-Layer Perceptron (MLP) with Back Propagation Learning to diagnose Parkinson's disease effectively with reduced number of attributes. Information Gain of all attributes is used as a measure to reduce the number of attributes. Mehdi et al. utilized data mining approaches for Genome-wide Association of Mood Disorders [18]. Six classifiers namely Bayesian Network, Support Vector Machine, Logistic Regression, Random Forest, Radial Basis Function and Polygenic Scoring Approach were compared and identified that simple polygenic score classifier performed better than others.

Filip and Caban [19] developed a Neural Network (NN) Model with an accuracy of 82.35% for predicting the likelihood of developing psychological conditions such as anxiety, behavioural disorders, depression and post-traumatic stress disorders. Shaikh, compared the performance of Artificial Neural Networks, Decision Tree and Naïve Bayes in predicting Parkinson's disease and Primary Tumour Disease and found that the accuracy is high in ANN for predicting Parkinson's disease and Naïve Bayes for Primary Tumour Disease [20].

The literature review shows that, a number of research works are going on in computerizing the diagnosis of mental health disorders and efforts are also taken to diagnose the mental health problems using different machine learning techniques in an efficient way. Various combinations of machine learning techniques (Hybrid techniques) are being employed to improve the accuracy of diagnosis with reduced set of features from profiles of patients. This research is to propose a Fuzzy Clustering based Bayesian framework for predicting the possibility of primary MHP like Attention Problems, Anxiety Symptoms, and Autism symptoms, Developmental Delay, ADHD and PDD among children. Early diagnosis of these types of problems among children allows early treatment and improves their quality of life.

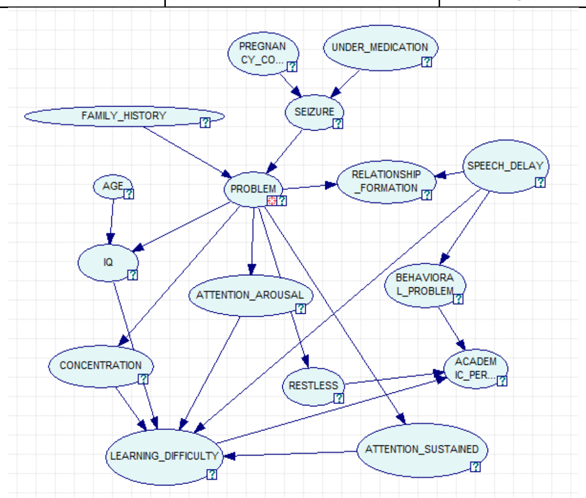
3. DATASET AND METHODOLOGY

The dataset for the research was collected from a clinical psychologist. It consists of 110 children profiles in the form of text documents. From the documents, 25 attributes relevant for the research were identified and values of those attributes were stored. The attributes identified from the dataset are listed in Table.1 in the Appendix at the end of the article. Pre-processing like identifying missing values and converting numeric values into nominal values, etc. were performed on the dataset. Feature selection was performed using Best First Search technique to identify more relevant attributes and eliminated irrelevant and redundant attributes. The more relevant attributes are shown in bold in Table.1.

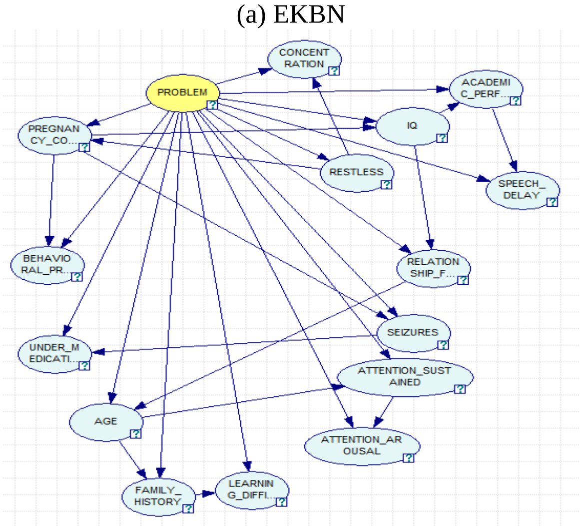
Table.1. Attributes identified from the Documents

Attribute	Meaning	Values
Age	Age group of the child, Infant / Early Childhood / Middle Childhood / Adolescent	{I, E, M, A}
Family History	Presence / Absence of psychological disorder in the family	{Y, N}
Pregnancy Complication	Presence / Absence of Complication during pregnancy	{Y, N}
Delayed Speech	Presence / Absence of delay in development of speech	{Y, N}
Under Medication	Whether the child is under any type of medicinal control	{Y, N}
Academic Performance	Whether the performance of the child is adequate/ inadequate in academics	{A, I}
Relationship Formation	Whether the child has adequate ability to socializ with peers, relatives and teachers	{A, I}
Behavioral Problem	Whether the child has any problem in behavior, Ex. Cheating, telling lies	{Y, N}
Concentration	Whether the child has adequate ability to focus hi attention on a particular object or activity	{A, I}
Restless	Whether the child is able to relax or involving in constant activity	{Y, N}
Seizures	Whether the child has a sudden surge of electrical activity in the brain which may cause unconsciousness, rigidness muscle stiffness, uncontrollable movements, etc.	{Y, N}
Learning Difficulty	Whether the child has any difficulty in acquiring age appropriate knowledge and skills	{Y, N}
Attention Aroused	Whether the child is psychologically alert, awake and attentive	{Y, N}
Attention Sustained	Whether the child is able to direct and focus cognitive activity on specific stimuli	{Y, N}
CBCL Score	A checklist filled by teacher/parent/self-report to identify problem behavior in children	{BC, EC, AC}
IQ Test Score	A standardized test to assess the intelligence leve of the child. The level may be Below Average, Average, Above Average	{BA, A, AA}
ADHD Positive	Screening test score for Attention-Deficit/Hyperactivity Disorder	{Y, N}
ODD Positive	Screening test score for Oppositional Defiant Disorder to measure argumentative and defiant behavior of the child	{Y, N}
Manic Episode Test Score	Screening test score for manic episode to check extreme elation or irritability in the mood of the child	{Y, N}
General Anxiety Disorder	Screening test score to measure the anxious level of the child	{Y, N}
Major Depressive Episode	Screening test score to measure the major depression level of the child	{Y, N}
CDI Score	Screening test score to measure the depressive symptoms of the child	{AC, EC, BC}
PDD Score	Screening test score to measure the level of	{MODERATE,NO, MILD}

	pervasive developmental disorder of the child	
Autism Score	Screening test score to identify whether the child has difficulty in communicating and forming relationships with other people and in using language and abstract concepts	{MODERATE,NO, MILD}
Problem	The class attribute specifying the mental health problem of the child	{ATT, ANX_SYM, PDD, DEV_DELAY, AUT_SYMP, ADHD}



(a) EKBN



(b) DBN

Fig.2. (a). Expert Knowledge based BN (EKBN) (b). Data based BN (DBN)

After pre-processing, 77 instances were obtained for training. On discussion with the psychologist, an Expert Knowledge-based Bayesian Network (EKBN) was constructed using GeNie tool and the parameters were learnt from the dataset. The model was validated with 2-fold cross-validation. The BN generated is shown in Fig.2(a). Using GeNie tool, a BN was generated automatically from data (DBN) and the generated network is shown in Fig.2(b). The predictive accuracies of EKBN and DBN are 48% and 58% respectively.

To improve the predictive accuracy of the BNs, Fuzzy Clustering based BN framework is proposed. The framework is detailed in section 4. This framework may be used to assist the professionals to diagnose the MHP effectively, if the known evidences of the patient are given as input to the framework.

4. PROPOSED FRAMEWORK

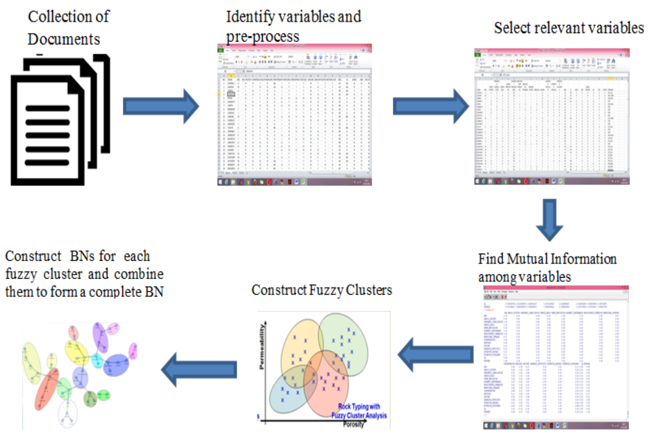


Fig.3. Fuzzy Clustering based Bayesian Framework

The proposed framework consists of the following phases. In the first phase, the required documents were collected from a clinical psychologist. Next, the documents have been pre-processed i.e. incomplete documents were removed, missing values were filled, numeric values have been categorized and stored in MS-EXCEL. Then, feature selection was performed on the dataset and the relevant features and their values were extracted. In the fourth phase, Fuzzy Clustering based Bayesian Network Structure learning Algorithm (FCBN) is applied to build the Bayesian network for predicting the basic MHP among children.

4.1 FUZZY CLUSTERING BASED BAYESIAN NETWORK STRUCTURE LEARNING FRAMEWORK (FCBN):

The framework is constituted of four steps:

Step 1: Find mutual information among nodes of BN.

Step 2: Based on mutual information matrix, group the nodes into various clusters using Fuzzy Clustering technique. This divides high dimensional structure space into several low dimensional structure spaces.

Step 3: Build local Bayesian network for each low dimensional space according to Bayesian Scoring function.

Step 4: Combine several local Bayesian networks based on fuzziness of nodes and build a complete Bayesian Network.

4.1.1 Find Mutual Information among Nodes of BN:

A Bayesian network consists of nodes and edges. The edges connecting the nodes represent the dependencies between them. The dependencies between the nodes have to be identified in order to learn Bayesian network. The mutual information between two nodes can tell us if the two nodes are dependent and if so, how close their relationship is. The mutual information between two nodes A and B is calculated using,

$$I(A,B)=\sum_{a,b}P(a,b)\log\frac{P(a,b)}{P(a)P(b)}$$

4.1.2 Group Nodes using Fuzzy Clustering Technique:

It is obvious that when the number of variables is large, the number of possible network structures grows exponentially. Robinson's recursive function helps us to determine the number of possible different network structures with *n* attributes. Hence, the number of nodes in a Bayesian network should not exceed 7 or 8 for decent time and search space. If a BN has more than 7 or 8 attributes, then this situation can be improved by dividing the high dimensional space into several low dimensional spaces. Low dimensional spaces can be identified by fuzzy clustering the nodes using the mutual information calculated in the previous step.

4.1.3 Construct Local Bayesian Networks for each Cluster:

Suppose, the fuzzy clustering technique divides *n* attributes into *k* clusters, then *k* Bayesian network structures are formed. The Bayesian structures are learnt based on Bayesian scores like Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) or Bayesian Dirichlet score (BDE) with two popular search techniques namely Tabu search and Hill Climbing search.

4.1.4 Combine Local Bayesian Networks to make a Complete Bayesian Network:

The local Bayesian networks are combined based on the fuzziness of nodes. For example, if a node '*x*' is found in Bayesian networks BN_1 and BN_2, combine the two networks and make it as one Bayesian network. Thus, the fuzzy nodes are identified to combine the networks and form a complete network.

4.2 IMPLEMENTATION AND EVALUATION

In this section, we describe the implementation of the proposed framework and its evaluation with different parameters. The algorithm was implemented using MS-Excel to store the data of mentally affected children, R tool to find mutual information among variables and form fuzzy clusters. GeNIe tool was used to construct and validate the network.

AGE	FAMILY_H	PREGNAN	SPEECH_D	UNDER_M	ACADEMI	RELATION	BEHAVIOF	CONCENT	RESTLESS	SEIZURE	LEARNING	ATTENTIO	ATTENTIO	IQ	PROBLEM
A	N	N	Y	N	BA	I	Y	I	Y	N	Y	Y	Y	AA	ATT
M	Y	Y	N	N	AA	A	Y	I	Y	N	N	Y	Y	AA	ANX_SYMP
A	Y	N	Y	N	A	I	Y	I	Y	N	Y	Y	Y	AA	ANX_SYMP
I	N	N	Y	N	A	I	Y	I	Y	N	N	N	N	AA	PDD
I	N	Y	Y	N	A	A	N	I	N	N	N	N	N	A	DEV_DELAY
I	Y	N	Y	N	A	A	Y	A	Y	N	N	N	N	A	PDD
I	N	N	N	N	A	A	N	A	N	Y	N	N	N	A	PDD
I	N	Y	Y	N	A	I	Y	A	Y	N	N	N	N	A	PDD
I	Y	Y	Y	N	A	I	Y	I	Y	Y	N	N	N	A	PDD
M	Y	N	Y	N	A	A	N	A	N	N	Y	Y	Y	BA	AUT_SYMP
I	Y	Y	Y	N	A	I	N	I	Y	N	N	N	N	A	PDD
I	Y	N	Y	N	A	I	N	I	Y	Y	Y	N	N	A	PDD
A	N	Y	N	N	BA	I	N	A	N	N	Y	Y	Y	BA	ATT
I	N	N	Y	N	A	I	N	A	N	N	N	N	N	A	PDD
E	N	N	Y	N	A	A	N	I	Y	N	N	Y	Y	A	PDD
I	N	Y	Y	N	A	I	N	A	N	N	N	Y	Y	A	DEV_DELAY
A	N	N	N	N	A	A	Y	I	Y	N	Y	Y	Y	A	ATT
A	N	N	N	N	A	A	N	I	N	N	Y	Y	Y	A	ATT
I	N	N	Y	N	A	A	N	A	N	N	N	Y	Y	A	PDD
M	N	N	N	N	A	I	Y	I	Y	N	N	N	N	AA	ATT
M	Y	N	Y	N	A	A	Y	I	Y	N	Y	Y	Y	AA	ATT
M	Y	N	N	N	AA	I	Y	I	Y	N	Y	Y	Y	AA	ATT
M	N	N	N	N	BA	A	Y	I	Y	N	Y	N	N	A	ATT
A	Y	Y	N	Y	BA	I	Y	I	Y	Y	Y	Y	Y	A	ATT

Fig.4. Screenshot of data stored in MS-Excel

The screenshot of data stored in MS-Excel is shown in Fig.4. The mutual information among the domain variables is shown in Fig.5. The Fuzzy K-Means clustering algorithm was executed in R tool and a sample membership degree matrix for 4 clusters is shown in Fig.6.

	AGE	FAMILY_HISTORY	PREGNANCY_COMPLICATION	SPEECH_DELAY	UNDER_MEDICATION	ACADEMIC_PERFORMANCE	RELATIONSHIP_FORMATION	BEHAVIORAL_PROBLEM
AGE	1.18	0.00	0.03	0.18	0.06	0.14	0.02	0.07
FAMILY_HISTORY	0.00	0.61	0.04	0.01	0.00	0.01	0.02	0.00
PREGNANCY_COMPLICATION	0.03	0.04	0.60	0.01	0.03	0.00	0.01	0.00
SPEECH_DELAY	0.18	0.01	0.01	0.68	0.00	0.07	0.03	0.00
UNDER_MEDICATION	0.06	0.00	0.03	0.00	0.39	0.02	0.00	0.00
ACADEMIC_PERFORMANCE	0.14	0.01	0.00	0.07	0.02	0.91	0.00	0.01
RELATIONSHIP_FORMATION	0.02	0.02	0.01	0.03	0.00	0.00	0.49	0.04
BEHAVIORAL_PROBLEM	0.07	0.00	0.00	0.00	0.00	0.01	0.04	0.61
CONCENTRATION	0.05	0.00	0.01	0.00	0.00	0.07	0.01	0.03
RESTLESS	0.04	0.00	0.02	0.00	0.00	0.00	0.00	0.06
SEIZURE	0.01	0.02	0.08	0.00	0.06	0.00	0.00	0.01
LEARNING_DIFFICULTY	0.10	0.00	0.00	0.00	0.00	0.04	0.00	0.01
ATTENTION_AROUSAL	0.11	0.00	0.00	0.03	0.00	0.03	0.01	0.00
ATTENTION_SUSTAINED	0.11	0.00	0.00	0.02	0.00	0.04	0.02	0.00
IQ	0.09	0.02	0.02	0.01	0.00	0.12	0.01	0.00
PROBLEM	0.49	0.04	0.03	0.18	0.01	0.14	0.02	0.10

	CONCENTRATION	RESTLESS	SEIZURE	LEARNING_DIFFICULTY	ATTENTION_AROUSAL	ATTENTION_SUSTAINED	IQ	PROBLEM
AGE	0.05	0.04	0.01	0.10	0.11	0.11	0.09	0.49
FAMILY_HISTORY	0.00	0.00	0.02	0.00	0.00	0.00	0.02	0.04
PREGNANCY_COMPLICATION	0.01	0.02	0.08	0.00	0.00	0.00	0.02	0.03
SPEECH_DELAY	0.00	0.00	0.00	0.00	0.03	0.02	0.01	0.18
UNDER_MEDICATION	0.00	0.00	0.06	0.00	0.00	0.00	0.00	0.01
ACADEMIC_PERFORMANCE	0.07	0.00	0.00	0.04	0.03	0.04	0.12	0.14
RELATIONSHIP_FORMATION	0.01	0.00	0.00	0.00	0.01	0.02	0.01	0.02
BEHAVIORAL_PROBLEM	0.03	0.06	0.01	0.01	0.00	0.00	0.00	0.10
CONCENTRATION	0.59	0.11	0.00	0.04	0.01	0.02	0.01	0.05
RESTLESS	0.11	0.59	0.00	0.01	0.02	0.05	0.01	0.07
SEIZURE	0.00	0.00	0.49	0.00	0.01	0.00	0.02	0.01
LEARNING_DIFFICULTY	0.04	0.01	0.00	0.67	0.01	0.00	0.04	0.26
ATTENTION_AROUSAL	0.01	0.02	0.01	0.01	0.61	0.34	0.02	0.15
ATTENTION_SUSTAINED	0.02	0.05	0.00	0.00	0.34	0.46	0.05	0.19
IQ	0.01	0.01	0.02	0.04	0.02	0.05	0.93	0.20
PROBLEM	0.05	0.07	0.01	0.26	0.15	0.19	0.20	1.44

Fig.5. Screenshot of Mutual Information Matrix

Membership degree matrix

	Clus 1	Clus 2	Clus 3	Clus 4
AGE	0.25	0.26	0.23	0.25
FAMILY_HISTORY	0.24	0.26	0.25	0.25
PREGNANCY_COMPLICATION	0.24	0.24	0.26	0.26
SPEECH_DELAY	0.25	0.26	0.24	0.25
UNDER_MEDICATION	0.24	0.25	0.26	0.25
ACADEMIC_PERFORMANCE	0.29	0.24	0.23	0.24
RELATIONSHIP_FORMATION	0.25	0.25	0.25	0.24
BEHAVIORAL_PROBLEM	0.26	0.24	0.25	0.24
CONCENTRATION	0.25	0.24	0.26	0.24
RESTLESS	0.24	0.24	0.27	0.25
SEIZURE	0.23	0.24	0.27	0.26
LEARNING_DIFFICULTY	0.25	0.25	0.25	0.25
ATTENTION_AROUSAL	0.24	0.26	0.25	0.26
ATTENTION_SUSTAINED	0.24	0.26	0.25	0.26
IQ	0.27	0.25	0.24	0.25
PROBLEM	0.25	0.26	0.23	0.25

Fig.6. Membership Degree Matrix for 4 Clusters

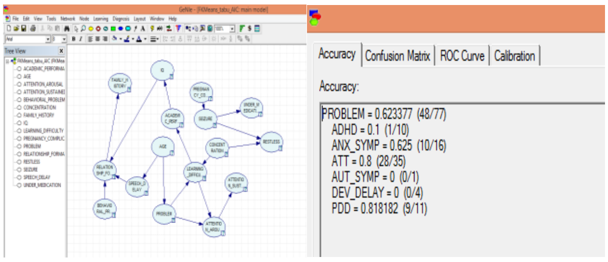


Fig.7. Screenshot of BN created and the results generated in GeNIe tool

For each cluster, the Bayesian networks were constructed in R tool using different fuzzy clustering algorithms and different Bayesian network scores. The Bayesian networks for each cluster are combined using the fuzzy node and a complete BN was generated. The complete BN was constructed in GeNIe tool to validate the results. The BN constructed and the prediction results of the network obtained using GeNIe tool are shown in Fig.7.

The performance of the proposed FCBN framework has been evaluated through several experiments with different number of clusters, different Fuzzy clustering methods and different Bayesian network scores. As the optimal number of clusters for a BN should be less than 7 or 8, we have used three numbers of clusters 2/4/6. The three Fuzzy Clustering methods used are Fuzzy K-Means, Fuzzy K-Medoids, Fuzzy K-Means with Entropy Regularization. Two popular search techniques namely, Tabu search and Hill Climbing search and three score types namely, Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and Bayesian Dirichlet Score (BDE) were considered to generate the best BN. The comparison of the results is shown in Table.2.

Table.2. Experimental Results of the three Fuzzy Clustering Algorithms

Clustering Method	**Search Type**	**Score Type**	**Number of Clusters**		
			2	**4**	**6**
Fuzzy K-Means	Tabu	AIC	0.68	0.62	0.60
		BIC	0.65	0.61	0.56
		BDE	0.68	0.61	0.61
	Hill Climbing	AIC	0.57	**0.66**	0.61
		BIC	0.66	0.53	0.49
		BDE	**0.69**	0.48	0.58
Fuzzy K-Medoids	Tabu	AIC	0.56	0.45	**0.66**
		BIC	0.48	0.45	0.65
		BDE	0.58	0.45	0.65
	Hill Climbing	AIC	0.45	0.60	**0.66**
		BIC	0.53	**0.66**	0.65
		BDE	0.53	0.52	0.65
Fuzzy K-Means with Entropy Regularization	Tabu	AIC	0.68	0.61	**0.66**
		BIC	0.65	0.57	0.64
		BDE	0.68	0.60	0.62
	Hill Climbing	AIC	0.57	0.61	0.64
		BIC	0.66	0.64	0.61
		BDE	**0.69**	0.60	0.64

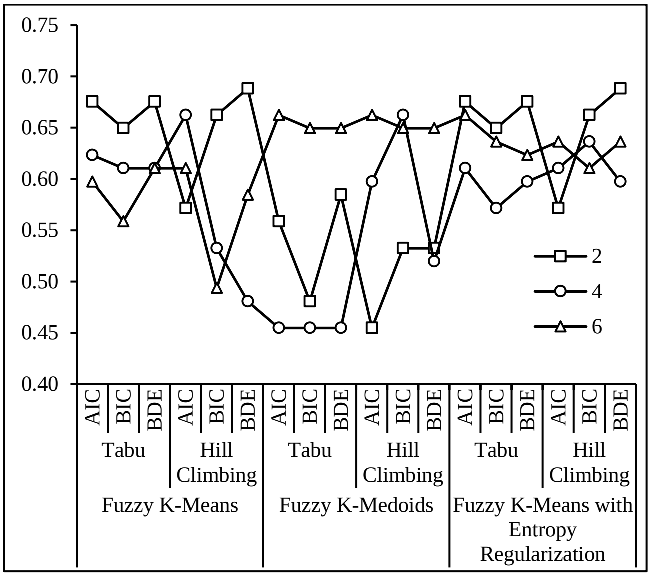


Fig.8. Comparison of Predictive Accuracy of the three Fuzzy Clustering Algorithms

The Table.2 shows that the Fuzzy K-Means and Fuzzy K-Means with Entropy Regularization performed better than Fuzzy K-Medoids when Hill Climbing search was employed with BDE Score. The graphical representation of Table.2 is shown in Fig.8. The graph shows that, when the number of clusters is less, the accuracy is more. And the Fuzzy K-Means and Fuzzy K-Means with Entropy Regularization algorithms performed better than the Fuzzy K-Medoids algorithm. Hill climbing algorithm performed well in searching the Bayesian Networks with BDE Score.

The comparison of predictive accuracies of EKBN, DBN and Fuzzy Clustering based BNs shows that the latter one (FCBN) is more accurate than the others.

5. CONCLUSION AND FUTURE WORK

Fuzzy Clustering based Bayesian Framework was proposed to diagnose the mental health problems of children effectively. Fuzzy clustering has been used to cluster the random variables based on mutual information between them. The clustered variables are then combined to learn smaller BNs. After generating BNs for all clusters, they are combined to form a complete network for prediction. An experimental analysis was performed over three algorithms, namely, Fuzzy K-Means, Fuzzy K-Medoids and Fuzzy K-Means with Entropy Regularization. Tabu Search and Hill Climbing Search techniques were employed to search the BNs on various scores like Akaike Information criterion, Bayesian Information Critierion and Bayesian Dirichlet Score. Fuzzy K-Means and Fuzzy K-means with Entropy Regularization with Hill Climbing and BDE Score performed well and the predictive accuracy are higher than others. In future, this framework can be modified and the predictive accuracy can be improved by altering the parameters of the algorithms. The number of variables taken into account may also be increased and the efficiency of the framework has to be calculated with other datasets.

ACKNOWLEDGMENT

We acknowledge Dr. Sangeetha Madhu, Clinical Psychologist, Chennai Institute of Learning and Development (CHILD), Chennai for her support in providing databases for the analysis and in providing expert knowledge.

REFERENCES

[1] Mental Health in India: A Wake-up Call, Available at: http://blogs.fortishealthcare.com/mental-health-india-wake-up-call/

[2] Bruce G. Buchanan and Edward H. Shortliffe, "*Rule-based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*", Addison-Wesley, 1984.

[3] Randolph A. Miller, Harry E. Pople and Jack D. Myers, "Internist-I, An Experimental Computer-based Diagnostic Consultant for General Internal Medicine", *New England Journal of Medicine*, Vol. 307, pp. 468-476, 1982.

[4] R.H. Yap and D.M. Clarke, "An Expert System for Psychiatric Diagnosis using the DSM-III-R, DSM-IV and ICD-10 Classifications", *Proceedings of the AMIA Annual Fall Symposium*, pp. 229-233, 1996.

[5] Kuryati Kipli, Abbas Z. Kouzani and Isredza Rahmi A. Hamid, "Investigating Machine Learning Techniques for Detection of Depression using Structural MRI Volumetric Features", *International Journal of Bioscience, Biochemistry and Bioinformatics*, Vol. 3, No. 5, pp. 444-448, 2013.

[6] D.H. Wolpert and W.G. Macready, "No Free Lunch Theorems for Search", *IEEE Transactions on Evolutionary Computation*, Vol. 1, No. 1, pp. 67-82, 1997.

[7] Rozita Yati Masri and Hajar Mat Jani, "Employing Artificial Intelligence Techniques in Mental Health Diagnostic Expert System", *Proceedings of International Conference on Computer and Information Science*, Vol. 1, pp. 495-499, 2012.

[8] David D. Luxton, "Artificial Intelligence in Psychological Practice: Current and Future Applications and Implications", *Professional Psychology: Research and Practice*, Vol. 45, No. 5, pp. 332-339, 2014.

[9] D. Razzouk, J.J. Mari, I. Shirakawa, J. Wainer and D. Sigulem, "Decision Support System for the Diagnosis of Schizophrenia Disorders, *Brazilian Journal of Medical and Biological Research*, Vol. 39, No. 1, pp. 119-128, 2006.

[10] Subhagata Chattopadhyay, Preetisha Kaur, Fethi Rabhi and U. Rajendra Acharya, "An Automated System to Diagnose the Severity of Adult Depression", *Proceedings of 2nd International Conference on Emerging Applications of Information Technology*, pp. 121-124, 2011.

[11] Luciano Comin Nunes, Placido Rogerio Pinheiro, Tarcisio Pequeno Cavalcante and Mirian Caliope Dantas Pinheiro, "Handling Diagnosis of Schizophrenia by a Hybrid Method", *Computational and Mathematical Methods in Medicine*, Vol. 2015, pp. 1-13, 2015.

[12] S.R. Basavappa, S.L.Rao, B. Harish, Shantala R Basavappa and Shobini L Rao, "Expert System for Dementia/ Depression Diagnosis", *Nimhans Journal*, Vol. 14, No. 2, pp. 99-106, 1996.

[13] Rashedur M. Rahman and Farhana Afroz, "Comparison of Various Classification Techniques using Different Data Mining Tools for Diabetes Diagnosis", *Journal of Software Engineering and Applications*, Vol. 6, No. 3, pp. 85-97, 2013.

[14] David Gil and B. Mangus Johnson, "Diagnosing Parkinson by using Artificial Neural Networks and Support Vector Machines", *Global Journal of Computer Science and Technology*, Vol. 9, No. 4, pp. 63-71, 2009.

[15] Jerzy Gomuła, Krzysztof Pancerz and Jarosław Szkoła, "Classification of MMPI Profiles of Patients with Mental Disorders-Experiments with Attribute Reduction and Extension", *Proceedings of International Conference on Rough Sets and Knowledge Technology*, pp. 411-418, 2010.

[16] Flavio Luiz Seixas et.al., "A Bayesian Network Decision Model for Supporting the Diagnosis of Dementia, Alzheimer′s Disease and Mild Cognitive Impairment", *Computers in Biology and Medicine*, Vol. 51, pp. 140-158, 2014.

[17] Anchana Khemphila and Veera Boonjing, "Parkinsons Disease Classification using Neural Network and Feature Selection", *World Academy of Science, Engineering and Technology*, Vol. 6, No. 4, pp. 15-18, 2012.

[18] Mehdi Pirooznia et al., "Data Mining Approaches for Genome-Wide Association of Mood Disorders", *Psychiatric Genetics*, Vol. 22, No. 2, pp. 55-61, 2012.

[19] Filip Dabek and Jesus J. Caban, "A Neural Network Based Model for Predicting Psychological Conditions", *Proceedings of International Conference on Brain Informatics and Health*, pp. 252-261, 2015.

[20] Tawseef Ayoub Shaikh, "A Prototype of Parkinson's and Primary Tumor Disease Prediction using Data Mining Techniques", *International Journal of Engineering Science Invention*, Vol. 3, No. 4, pp. 23-28, 2014.