International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-3, September 2019

Improved Reasoning with Uncertainty Based Significant Feature Subset Selection for Alzheimer's Disease Detection

K.Yemunarane, A.Hema

Abstract*: *In this modern era neurodegenerative disorder of undefined causes affects the older adults and it becomes most cause of dementia. The Alzheimer's disease is one of such neurodegenerative disorder which is very complex and hard to predict in the early stage. With evolving advancement in the field of machine learning, it is possible to predict the early stage of AD and diagnosing in initial stages may produce effect result for their further quality and healthy life. But uncertainty in determination of Alzheimer's is a toughest challenge for the researchers in the field of machine learning. This paper aims to overcome the uncertainty in discovering dementia and non-dementia victims of Alzheimer's by devising an improved reasoning with uncertainty based prominent feature subset selection using modified fuzzy dempster shafer theory (IRU-DST). For Alzheimer's disease prediction the dataset is used form OASIS dataset. The performance of the proposed IRU-DST is validated using fuzzy artificial neural network. The simulation results proved the performance of the IRU-DST achieves better results comparing the other sate of arts, by gaining high accuracy rate and it also minimize the error rate considerably with the ability of handling uncertainty.

Keywords* : *Neurodegenerative disorder, Alzheimer's disease, Machine learning, Fuzzy logic, Dempster shafer theory, Fuzzy Artificial Neural Network

I. INTRODUCTION

Around 44 million of people are suffering from Dementia, it is a term which describes variant brain disorder. It affects thinking, memory, emotion and behaviour [1]. The dementia is classified into different kinds, major source of dementia is Alzheimer [2]. It is a kind of permanent syndrome of brain which consistently affects both behaviour and cognition, which consequences in impairment in the capability to do their routine activities. This type of dementia, affects 6 percentage of aged peoples over 65. Prediction of Alzheimer's (AD) in earlier stages is a major challenge for healthcare in the 21st century [3]. The signs of Alzheimer's are poor judgment, misplacing things, abnormal moods, decision making, verbal communication ,impairment of movements, memory loss. Under three diverse stages, AD diagnosis process is done like neuro psychological assessment, general physician and MRI scan [4]. Hence, diagnosing the Alzheimer's in earlier stages may assist the patients to either control it or slow down its aggressiveness and it leads to have quality life for their rest of the lifetime.

The evolution of machine learning, data mining and artificial intelligence greatly influence the process of prediction, in which they involved in mining hidden information that may help the experts to take right decision as earlier as possible. This paper aims on developing an uncertainty handling problem by introducing modified fuzzy dempster shafer theory which performs feature selection based on evidence. The feature selection is also an important stage which improves the quality of dataset more prominently for achieving better prediction results.

II. LITERATURE REVIEW

This section discusses about the various existing works on Alzheimer's disease detection.

David et al. [5] reported that there is a worldwide rise of 4.6 million new dementia cases each year. If no mortality changes or any other effective strategies to prevent the Alzheimer's is not considered then this will double every 20 years.

Sandhya Joshi et.al [6] in their work used different machine learning models like multilayer perceptron , decision tree, bagging, genetic algorithm to classify and diagnosis Alzheimer's disease.

Devi Parikh et.al [7] devised a novel fused method to diagnose Alzheimer using two different data sources that contains complementary information. The supervised learning model is used for classifying the Alzheimer's disease at very stages.

Claudia Plant et.al [8] designed an automated MRI based bran atrophy patterns of Alzheimer's disease using support vector machine, voting feature interval and bayes statistics. To derive a quantitative pattern matching scheme they used voting feature interval which predicts the possibility of conversion from MCI to AD in future.

Stefan Kloppel et.al [9] developed a classification model which receives MR scans as input to detect the presence of Alzheimer. grey matter of MRI scan is identified and categorized by adapting linear support vector machine.

Ali Hamouet et al [10] anticipated a clustering approach using MRI scans to detect the Alzheimer's disease.

Manuscript published on 30 September 2019
* Correspondence Author
K.Yemunarane*, Computer Science, Kongunadu Arts and Science College, Email: yemunaranekumaravel21@gmail.com
Dr.A.Hema, Computer Science, Kongunadu Arts and Science College, Email:hemasrgm@yahoo.com

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC-BY-NC-ND license http://creativecommons.org/licenses/by-nc-nd/4.0/

They used decision tree refinement scheme for analyzing and clustering the dataset. They discovered the influencing variants which greatly contributes in decision making.

J. Shane Kippenhan et.al [11] devised a neural network classification model to discover Alzheimer's presence on PET images with high and low resolutions. They trained the network to differentiate among both normal and abnormal cases

Obi J.C1.and Imainvan A.A [12] analyzed the standard procedure applied by the experts to diagnose AD. The statistical methods are applied based on their demographic information and MRI scan information and sociological behaviors of the victims.

Sandhya Joshi et.al [13] developed a classification model which involves in disorder related to neuro degenerative such as vascular, Parkinson and Alzheimer's disease based on the general risk factor. They used neural networks, decision tree, random forest and Radial basis function for classification.

Javier Escudero et.al [14] introduced machine learning approaches which infer the hidden knowledge about Alzheimer's disease using various classification cost effective models.

Roman Filipovych et.al [15] developed a supervised classification model that works on MR images by applying support vector machine. This work performs the preprocessing, feature extraction and classification of different stages of Alzheimer's.

Even though, there are lots of existing works available in Alzheimer's prediction system, there is no proper proof for handing inconsistencies and vagueness in the diagnosis of Alzheimer's. Henceforth, the ultimate objective of this paper is to handle the uncertainty by developing feature selection using evidential approach for effective Alzheimer's disease prediction.

III. METHODOLGY

A. Dempster Shafer Theory

Dempster Shafer theory (DST) is an evidence based model along with generalization of Bayesian, which handles imprecise probability and uncertainty theories. Belief function are mass is used to determine possible outcomes and their proportions are denoted as interval values which ranges in between belief and possibility.

$$\text{belief} \leq \text{plausibility}.$$

Belief is defined as a all subsets of masses summed value of a given hypothesis set. Either a particular or hypothesis event's support is directly estimated using belief value. The strength of evidence of a proposition, is measured using belief in term of favoritism. The value ranges from 0 to 1 which signifies evidence to certainty. one minus sum of all sets masses whose intersection with hypothesis empty is denoted as plausibility.

The Frame of Discernment (Θ)

All the sets in the whole space i.e exhaustive space is described as hypothesis space. Θ symbol represents that the elements and the frame are mutually exclusive. If there are n numbers of elements in a set then the power set consist of 2^n elements. For example, if: n = {a,b} then 2^n = { ∅, {a},{b},n}

BPA (Basic Probability Assignment)

The basic probability assignment represents assignment of belief mass. Every Belief mass is allotted to a subset. In this all the subset is known as power set. For the given set 0 and 1 is assigned. Likewise, the probability value persists.

If Θ is the frame of discernment, then a function m: $2^n \rightarrow$ [0, 1] is termed as Basic Probability Assignment with two different properties

- The mass value of an empty set is always zero, which is signified as

$$\text{Mass}(\emptyset) = 0$$

- The masses of the remaining power set members are sum to a total value of 1 and it is represented as

$$\sum_{X \in 2^n} Mass(X) = 1$$

Through the assignment of mass, probability interval's namely upper and lower limits can be well-define. Thus, it converts the interval of probability in to a precise set of interest. It is achieved by two important measure known as belief and plausibility.

Belief (Bel)

The Belief measure which is signified as bel(X) for a given set X, belief is calculated by summing all the mass values of subsets of interest. The frame of discernment Θ and a body of empirical evidence {m(B1), m(B2), m(B3)….}, the belief dedicated to X ε Θ is represented as

$$\text{Bel}(X) = \sum_{Y|Y \subseteq X} Mass(Y) = 1$$

Plausibility Function (Pl)

The plausibility pl(X) is the summing all mass value of set Y which intersects the set of attention of X. This is signified as follows:

$$\text{Pl}(X) = \sum_{Y|Y \cap X \neq \emptyset} Mass(Y)$$

Belief Range

The interval [Bel (A), Pl(A)] is called belief range. Plausibility (Pl) and Belief (Bel) are associated as follows

$$\text{Pl}(X) = 1 - \text{Bel}(\bar{X})$$

On contrary wise, the finite X, whose subset Y's belief measure is known, then it is possible to discover the masses of mass(X) using the following function

$$\text{Mass}(X) = \sum_{Y|Y \subseteq X} (-1)^{|X-Y|} \text{Bel}(Y)$$

Where |X-Y| is the variance of two sets cardinalities. In this standard dempster shafer theory the mass (probabilities of events) is not well defined if the set used in infinite, so this proposed work defines the mass value using Poission distribution.

$$\text{Poission}(x: \mu) = (e^{-\mu})(\mu^x) / x!$$

Where e is the Euler's value, μ average value of the historical available value and X is the predicted average value.

B. Dempster's Combination Rule

In case of multiple independent evidences are available, using dempster rule of combination, the combined evidence can be computed.

$$mass_{1,2}(\emptyset) = 0$$

$$mass_{1,2}(X) = (mass_1 \oplus mass_2)(X) = \frac{1}{1-H} \sum_{Y \cap Z = X \neq \emptyset} mass_1(Y) mass_2(Z)$$

Where $H = \sum_{Y \cap Z = \emptyset} mass_1(Y) mass_2(Z)$ is a metric for determining amount of conflict among two mass sets Y and Z. The main disadvantage in this generalized dempster rule of combination is that, the normalization factor, H-1 influences on entirely disregarding conflict and passing on any likelihood mass associated with controversy to null set, which leads to counterintuitive results.

C. Improved reasoning with uncertainty based prominent feature subset selection using modified fuzzy dempster shafer theory (IRU-DST)

This proposed IRU-DST collects Alzheimer's dataset from OASIS repository. The dataset is initially normalized to make all the attributes to fall under same range of value 0 to 1. Then most significant feature subset is evaluated using IRU-DST. With the resultant features subset its optimality is validated by using fuzzy ANN classifier. The detailed description of the each stage is explained in the following subsections the complete architecture of IRU-DST is shown in the figure 1.

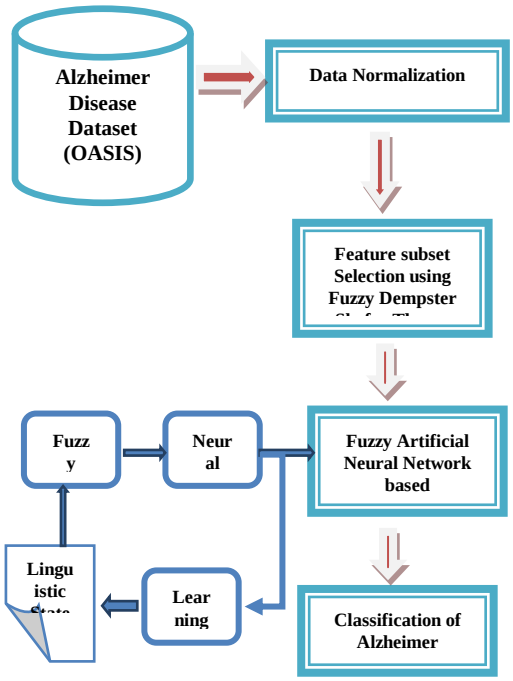

Figure 1: Overall Architecture of IRU-DST for Alzheimer's Disease Detection

Dataset Description

The team has found MRI related data that was generated by the Open Access Series of Imaging Studies (OASIS) project that is available both, on their website and Kaggle that can be utilized for the purpose of training various machine learning models to identify patients with mild to moderate dementia. OASIS dataset contains 416 images from infants to 18 years old ages. In this research work totally 373 instances with 15 attributes in which 12 attributes along with one class label named as group and remaining two attributes are identifiers comprised in longitudinal dataset. The below table describes the dataset description of both longitudinal Information of Alzheimer's Disease Database provided by OASIS [16].

Table 1: Longitudinal dataset Attributes

Features	Description
Subject ID	Subject Identification
MRI ID	MRI identification number
Visit	No. Of times Visited
MR Delay	MR Delay time Contrast
Sex	Gender (Male or Female)
Hand	Dominant Hand
Age	Age in Years
EDUC	Years of education
SES	Socioeconomic Status
MMSE	Mini Mental State Examination
CDR	Clinical Dementia Rating
eTIV	Estimated Total Intracranial Volume
nWBV	Normalize Whole Brain Volume
ASF	Atlas Scaling Factor
Group	Demented, Non Demented

D. Feature Selection Using Fuzzy Dempster Shafer Theory

This section discusses about the proposed model of fuzzy dempster shafer theory for feature selection. Initially the value of the dataset is converted to fuzzy representation, then mass value of each instances of the dataset is calculated and using the combined dempster shafer rule, uncertainty existence among feature selection is handled and significant attributes are chosen.

Fuzzy Representation

The dataset of Alzheimer's Disease is represented in the crisp value's, to handle the problem of uncertainty, these crisp values are converted to fuzzy values after performing normalization process using minmax method.

Apply normalization process on the dataset

$$\text{Normalized value} = \left[\frac{\text{x - Minimum Value}}{\text{Maximum Value - Minimum Value}}\right]$$

Then the fuzzy membership value for each attributed in the Alzheimer's dataset is computed using the following formula

$$\mu_A(x) = \begin{cases} 0 & (x < a) or (x < d) \\ \dfrac{x-a}{b-a} & a \le x \le b \\ 1 & b \le x \le c \\ \dfrac{d-x}{d-c} & c \le x \le d \end{cases}$$

Figure 2: (a) Formula for membership representation

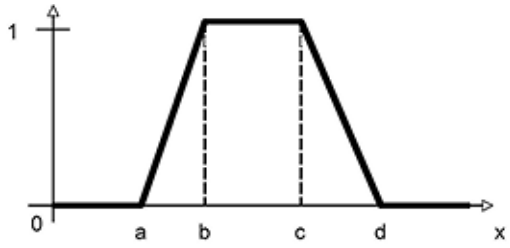

Figure 2: (b) pictorial representation of trapezoidal function

Where x is the value of the concern instances attribute, a and d represent feet of trapezoidal, b and c locate at shoulder of trapezoidal.

E. Modified Dempster Shafer Combine Rule

To overwhelm the issue of counterintuitive results, Yager's rule is implemented to handle the uncertainty. The ground truth mass assignment is used instead of basic probability mass assignment. The main different is reflected in the normalization parameter and the mass is attributed as the universal set.

$$q(A) = \sum_{B \cap C \subseteq A} m_1(B) m_2(C)$$

The intersection of B and C subsets are denoted as A and q(A) represents assignment of probability related to A. Normalization factor is not involved in this process. This is referred as Yager's rule of combination. Yet this method is not associative, the integrated structure q(A) allows to use any amount of evidence pieces. Let's consider bpa of n structure is denoted as $mass_1$, $mass_2$,…, $mass_n$. Let Foc_i signifies focal elements set related with the ith belief construction ($mass_i$) which are subclasses of the general setX. A_i signifies an element of the focal set. Then the amalgamation of n basic probability assignment assemblies is well-defined as follows:

$$q(A) = \sum_{\cap_{i=1}^{n} A_i = A} mass_1(A_1) mass_2(A_2) \ldots mass_n(A_n)$$

the integrated construction of q(A) can be altered depending on new evidence. This is achieved while ground probability assignment is combined with the new evidence along with existing mixture. Basic variance among ground probability assignment is absenteeism of the normalization factor (1-K). By adapting formula of yager's, the circumvents normalization is achieved by permitting ground probability mass of null set to be greater than 0, i.e.

$$q(\emptyset) \geq 0$$

Later Yager's augments conflict value q($\emptyset$) to ground probability assignment of the universal set $q(X)$, which yields the alteration of the ground probabilities to bpa of the universal set $m^Y(X)$:

$$m^Y(X) = q(X) + q(\emptyset)$$

F. Feature Selection using IRU-DST

The table 2 shows the sample example of the IRU-DST based feature selection on Alzheimer's disease detection. Here four different attributes are considered for observation. Based on the obtained Mass value the belief and the plausibility for two different classes are computed.

As for each feature in Alzheimer's dataset, the mass value is assigned and denoted as Mass(F_k), their corresponding belief and plausibility are determined based on their hypothesis stated in the powerset. The belief is computed by summing its own basic assignment of mass with those of all of its subset's

Example:

BL(($FS1_5$) = Mass ($FS1_2$) +Mass($FS1_3$)+Mass($FS1_5$) = 0.2+0+0.1 = 0.3

The plausibility is measure by summing all the basic assignment mass value which have got at least one hypothesis with those of the concern statement.

PL(($FS1_5$) = Mass ($FS1_2$) +Mass($FS1_3$)+Mass($FS1_4$) + Mass($FS1_5$) +Mass($FS1_6$) + Mass($FS1_7$)

= 0.2+0+0.5+0.1+0+0.1 = 0.9

Table 2: Sample Example of IRU-DST Based Feature Selection AD Detection

Demented			Nondemented			Powerset
Mass($FS1_k$)	Bel(D)	Pl(D)	Mass($FS2_k$)	Bel(ND)	Pl(ND)	
0.1	0.1	0.7	0.2	0.2	0.5	{h1}
0.2	0.2	0.9	0	0	0.6	{h2}
0	0	0.2	0.1	0.1	0.8	{h3}
0.5	0.8	1	0	0.2	0.8	{h1∪ h2}
0.1	0.3	0.9	0.4	0.5	0.8	{h2∪ h3}
0	0.1	0.8	0.1	0.4	1	{h1∪ h3}
0.1	1	1	0.2	1	1	{ h1∪ h2∪h3}

To perform optimized Alzheimer's disease detection, in this work modified fuzzy dempster shafer theory-based feature selection method is used to discover the most significant features, which contribute more in classification of Alzheimer's disease. The frame of discernment is {demented, Non-demented} with Non-demented refers a normal status that is absence of Alzheimer's and demented refers to presence of Alzheimer's disease (AD). In this work, AD dataset is validated using ten-fold cross validation which randomly partition the dataset into ten subsets of equal size and during each execution one subset of dataset is considered as test data and remaining nine subsets are involved in training process. The training process obtains the threshold value t to construct the mass function. The complete dataset size is 373, so the training data size will be 335 or 336. For each individual attribute, the probability function is applied with Poisson distribution. A general assumption is that lower-valued items tend to represent normal data, so the mass functions for each attribute can be modeled using sigmoid functions:

- $m(\text{non-dementia}) = (1+e^{(v-t)})^{-1}$
- $m(\text{dementia}) = 1-m(\text{non-dementia})$

where *v* is the value of the test data item for that attribute. The Alzheimer's datasets comprised of data items which are integer sand hence the functions consist of discrete values only. The mass functions are used to assign mass values to each attribute for each item in the test data, and these are then combined using Yager's Combination Rule to obtain overall normal and abnormal mass values. For a given data item, if the mass value of the abnormal hypothesis is larger than that of the normal hypothesis, then that item is classified as Alzheimer's, otherwise it is classified as no Alzheimer's. The architecture is represented schematically in Figure 3.

Classification accuracy *ACC* is used to assess the quality of the feature subset generated by modified Fuzzy Dempster Shafer Rule. This is given by:

$$Acc = \frac{Number\ of\ Correctly\ Classified\ instances}{Total\ of\ Number\ of\ instances}$$

where Nc is the number of correctly classified items and Nt is the total number of data instances. Classification is carried out using IRU-DST on all the attributes and powerset of number of input attributes.

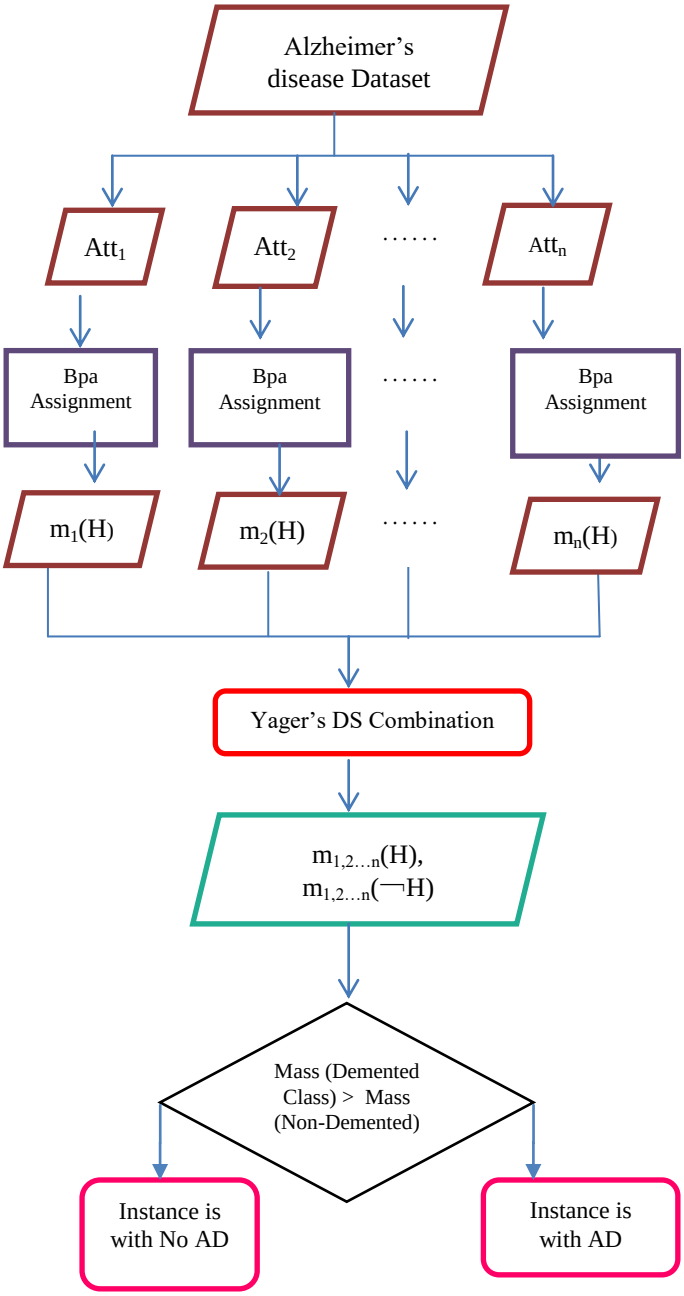

Figure 3: Work Flow of Yager's Dempster shafer combination Rule for feature subset selection

G. Fuzzy Artificial Neural Network

After performing feature selection using modified fuzzy Dempster shafer Theory (IRU-DST), to validate the performance of this model the fuzzy artificial neural network is used for determining the error rate of classification with different feature subset's generated by Principal Component Analysis (PCA) , Mutual Information (MI), Standard Dempster Shafer theory (DST)and Modified Fuzzy Dempster Shafer Theory with whole feature set. The Figure 3 illustrates the workflow of fuzzy artificial neural network classifier involving in Alzheimer's disease prediction.

Layer 1: This is the first layer which receives the input from the real valued input variable of Alzheimer's Dataset.

Layer 2: This second layer transformed layer represents the fuzzification value done using the formula

$$(x) = \begin{cases} 0 & (x < a) or (x < d) \\ \dfrac{x-a}{b-a} & a \le x \le b \\ 1 & b \le x \le c \\ \dfrac{x-a}{b-a} & c \le x \le d \end{cases}$$

Layer 3: This third layer is the hidden layer which calculates the weight value, bais value with the corresponding input values. The i-th node computes the ratio of the i-th rule's firing strength to the sum of all rules' firing strengths.

$$net_i^{(1)} = \sum_{j=1}^{n} w_{ij}^{(1)} X\, x_j,$$

Layer 4: The fourth layer represents output layer. Output layer is performed as follows:

$$net^{(2)} = \sum_{j=1}^{n^1} w_{1^i}^{(2)} \times a_i^{(1)}$$

$$\boldsymbol{a^{(2)} = f\,(net^{(2)})}$$

Layer 5: The fifth layer which is a circle node labeled Σ that sums all incoming.

Once the input patterns are passed to the network it is acquire by the input nodes. Here input nodes are passive that is it just receives the input values and pass it to the next layer. The next layer is the hidden layer once it receives the value it multiplies those values with the weight assigned to the concerned connection. According to the Eq. (1) the input to a single node is weighted as follows

$$network_j = \sum_i wt_{ij} \,.\, \text{Out}_i \qquad (1)$$

where wt_{ij} represents the weights among $node_i$ and $node_j$, and out_i is the output from node i such as Eq. (2).

$$\text{Out}_j = \phi(network_j) \quad (2)$$

Here ϕ is the activation function. In this the hidden layer and the output layer uses the non-linear activation function known as sigmoid which has S shaped curve applied to the weighted sum of input values prior to the signal processes to the subsequent layer. The derivative of the sigmoid function is shown in Eq. (3).

$$\phi'(net_j) = \phi(net_j)\,(1 - \phi(net_j)) \quad (3)$$

Finally, the resultant output obtained and the desired output is compared, if they are different their difference is assigned as error.

IV. EXPERIMENT RESULTS

The simulation of proposed modified fuzzy dempster shafer theory which acts as reasoning with uncertainty in feature subset selection of Alzheimer's disease detection is done using MATLAB software. The three existing feature subset selection approaches namely PCA, MI and DST are used for comparison with IRU-DST. The MRI related dataset of Alzheimer's disease is collected from Open Access Series of Imaging Studies (OASIS) [16] and also provided by Kaggle [17]. This dataset consist of 373 instances with 13 attributes with two attributes are record identifiers.

The table 3 below shows the feature subset selected by four different models. In which the PCA generated 10 features, MI chooses 8 features as their feature subsets, DST selects 6 attributes as feature subset. The proposed IRU-DST chooses 4 features contributes more in detection of dementia and non-dementia of Alzheimer's they are MR Delay, SES, CDR and eTIV.

Table 3: Feature subsets generated by four Different Feature Selection Methods

Feature selection Methods	Feature Selected	No. of Feature Selected
PCA	Visit, MR Delay, Gender, EDUC, SES, MMSE, CDR, eTIV, nWBV,ASF	10
MI	visit, MR Delay, SES, MMSE, CDR, eTIV, nWBV,ASF	8
DST	MR Delay, SES, MMSE, CDR, eTIV, nWBV	6
IRU-DST	MR Delay, SES, CDR, eTIV	4

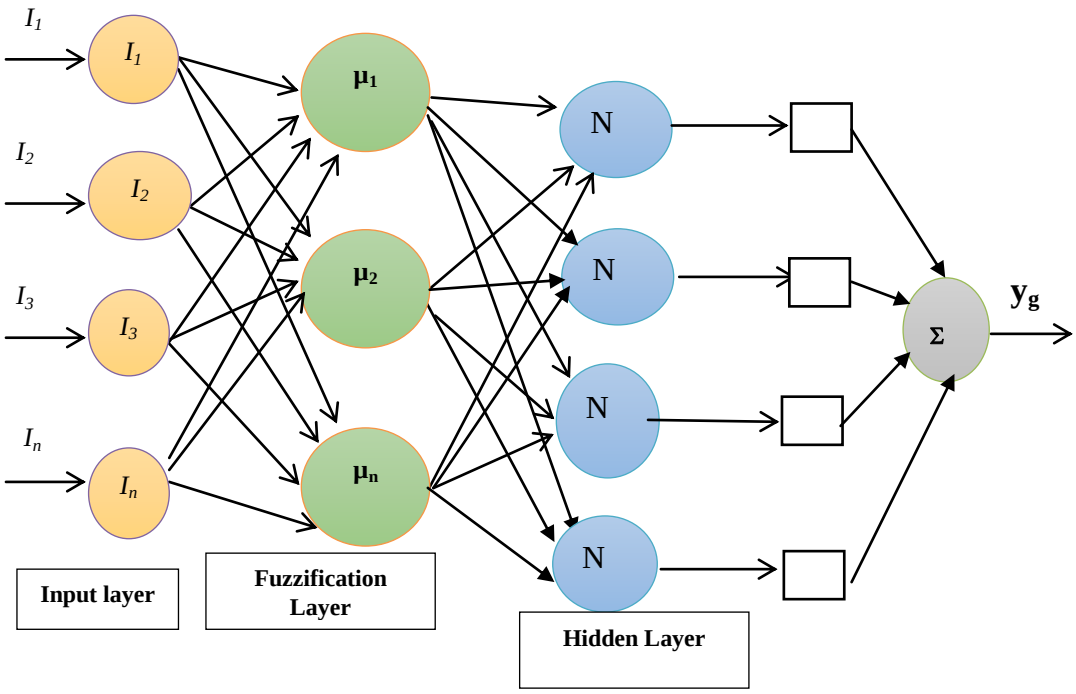

Figure 3: Layers of Fuzzy Artificial Neural Network

A. Evaluation Metrics

To determine the performance of the four different feature subset selection models three metrics are used. They are Precision, Recall and F-Measure

Precision: It is the measure of determining out of demented and non-demented, how much the classification model predicted correctly. It should be high as possible.

$$\text{Precision} = TP/(TP+FP)$$

Where TP represents when the actual class value is yes and the predicted value is also yes and the FP denotes when actual class value is no but the predicted class is yes.

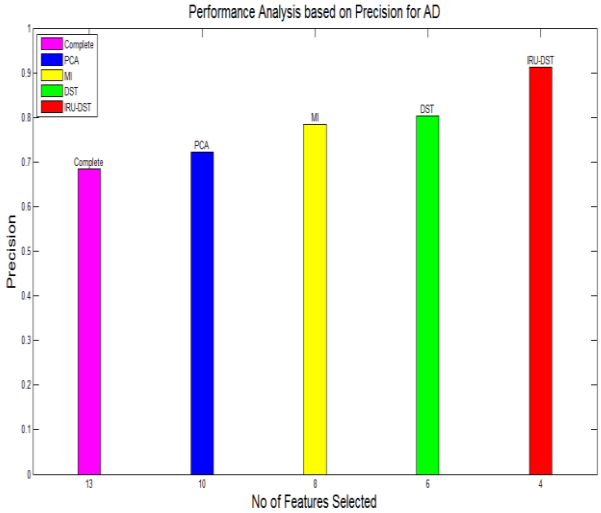

Figure 4: Performance Comparison based on Precision for Alzheimer's disease Detection using four different feature selections with fuzzy ANN as Classifier

Recall: It is the evaluation measure discovering out of all the positive classes, how much the classification model predicted correctly. It should be high as possible.

$$\text{Recall} = TP/(TP+FN)$$

Where FN refers to when actual class is yes but predicted class in no.

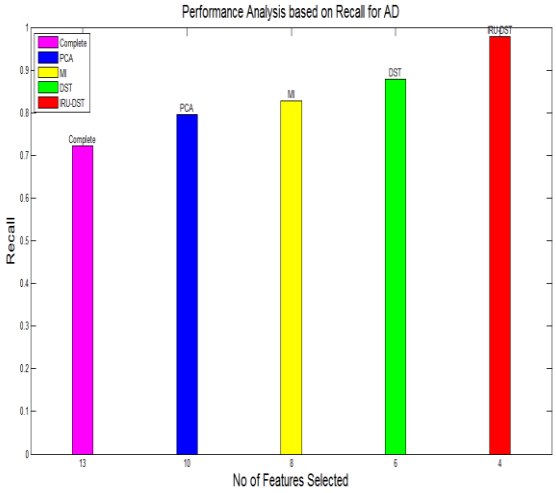

Figure 5: Performance Comparison based on Recall for Alzheimer's disease Detection using four different feature selections with fuzzy ANN as Classifier.

F-Measure: F-score helps to measure Recall and Precision at the same time. It is the weighted average of both the measure precision and recall.

$$F(CL_r, CS_i) = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} * \text{Precision}}$$

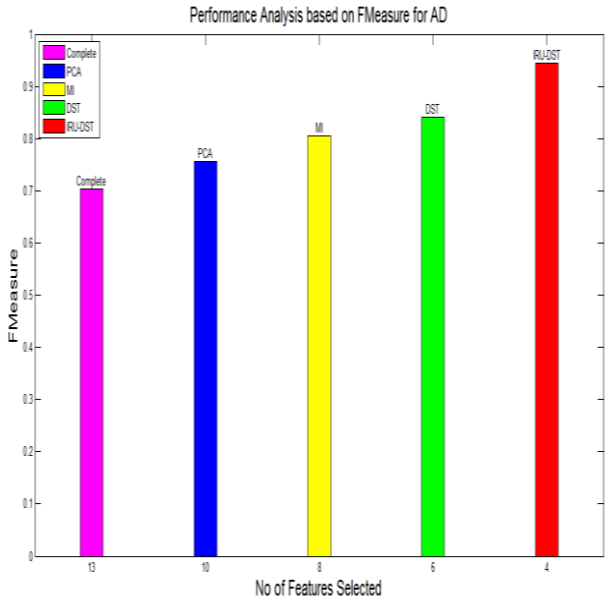

Figure 6: Performance Comparison based on F-Measure for Alzheimer's disease Detection using four different feature selections with fuzzy ANN as Classifier.

The results of the figure 6 show the performance of the modified fuzzy dempster shafer theory achieved better results on determining dementia and non-dementia patients for Alzheimer's disease detection. This will help in early detection of Alzheimer's and to advice the victims about its symptoms. The fuzzy ANN is used for training and testing the Alzheimer's disease dataset using four different sets of feature subsets generated by each feature selection approaches used in this work. The f-measure value of Fuzzy ANN produces higher results while using for IRU-DST based Alzheimer's prediction process. The ability to handle the uncertainty is as importance factor of proposed model IRU-DST.

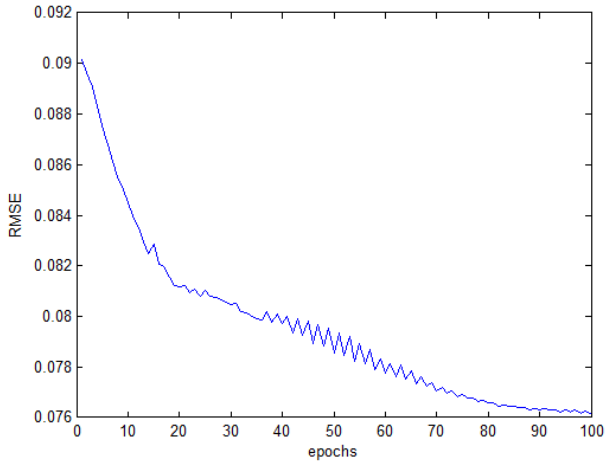

Figure 7: RMSE value ANN using Improved Fuzzy Dempster Shafer theory feature subset(IRU-DST)

During initial process of fuzzy ANN the error rate was high and as with use of backpropagation it adjust its weight values in the hidden layer and produces optimal results once it is trained for the sufficient iteration using the input of feature generated by modified fuzzy dempster shafer theory which is shown in the figure 7.

Table 4: Performance Analysis of Proposed IRU-DST based Feature subset Selection and Classification using Fuzzy ANN

Measures	Instances with %	
Correctly Classified Instances	361	96.78%
Incorrectly Classified Instances	12	3.22%

Table 4 shows the correctly classified instances of both demented and non-demented are 96.78% and the misclassified outcome is 3.22% obtained using the proposed model.

Table 5: Confusion Matrix for Alzheimer's disease prediction using Fuzzy ANN + IRU-D

	ND	D
ND	**187**	**3**
D	**9**	**174**

The table 5 depicts the correctly classified instances as 361 and the incorrectly classified as 9 out of 373 instances.

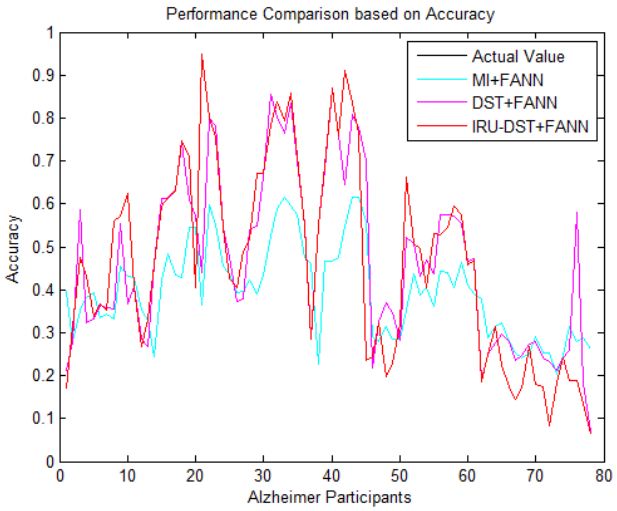

Figure 8: Performance Analysis based on Accuracy of Classification by FANN with three different feature subset selections

The figure 8 shows the ability of the fuzzy ANN using improved feature subset selection method (IRU-DST) increase the accuracy of classification of Alzheimer's disease detection.

The process of handling uncertainty in determining the significant feature subset is done by representing each features in the dataset to the degree of membership, clearly reveals the most contributing features which influences in classifying dementia and non-dementia of Alzheimer's. The other models works on probability basis and their feature selection are done on probability based without any evidence thus they fail to produce optimal accurate results.

V. CONCLUSION

This paper proves the eminent usefulness of IRU-DST theory to express uncertain judgments of experts in early stage detection of Alzheimer's disease. This proposed work IRU-DST, improves the quality of dataset by handling raw dataset by imputing missing values and it produces potential feature subset based on the evidence obtained by the belief and plausibility of each attribute involved in detection of Alzheimer's presence or absence. The issues of omitting contradiction occurrence in detection of Alzheimer's disease detection by standard DST are well handled in this IRU-DST. Thus the model IRU-DST developed in this work overcomes the counterintuitive results due to uncertainty more intelligently and increases the accuracy rate of classifying dementia and non-dementia of Alzheimer's in older adults and it also decreases the rate of false alarms more effectively.

REFERENCES

1. http://www.capitalfm.co.ke/lifestyle/2013/12/06/44-million-nowsuffer-from-dementia-worldwide.
2. Viswanathan A, Rocca WA, Tzourio C. Vascular risk factors and dementia: How to move forward? Neurology 72:Pp368–74, 2009
3. Alzheimer's Association. 2018 Alzheimer's disease Facts and Figures. Alzheimer's Dement 2018;14(3):367-429
4. Veitch, D.P., Weiner, M.W., Aisen, P.S., Beckett, L.A., Cairns, N.J., Green, R.C., Harvey, D., Jack, C.R., Jr., Jagust, W., Morris, J.C., Petersen, R.C., Saykin, A.J., Shaw, L.M., Toga, A.W., and Trojanowski, J.Q, understanding disease progression and improving Alzheimer's disease clinical trials: Recent highlights from the Alzheimer'sDisease Neuroimaging Initiative. Alzheimer's Dement 15, 106-152, (2019).
5. David P Salmon, Mark W. Bondi, Neuropsychological Assessment of Dementia, Access NIH public, PubMed central, US national library of medicine National Institutes of Health,60: 257–282, 2010.
6. Sandhya Joshi et.al "Classification and treatment of different stages of Alzheimer's disease using various machine learning methods", International Journal of Bioinformatics Research, Volume 2, Issue 1, pp-44-52, 2010
7. Devi Parikh, et.al "Ensemble Based Data Fusion for early Diagnosis of Alzheimer's Disease", Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference Shanghai, China, 1-4, September 2005
8. Claudia Plant et.al "Automated detection of brain atrophy patterns based on MRI for the prediction of Alzheimer's disease" Neuro Image 50 (2010) ,162–174 published by ELSEVIER
9. Stefan Kloppel et.al "Automatic classification of MRI Scans in Alzheimer's disease" published in Brain (2008), 13, PP 681-689 2008
10. Ali Hamou et.al "cluster analysis of MR Imaging in Alzheimer's disease using decision tree refinement" International journal of Artificial intelligence, Vol6, PP 1-10, , spring 2011
11. J.ShaneKippenhan et.al "Neural-Network Classification of Normal and Alzheimer'sdisease Subjects Using High-Resolution and Low-Resolution PET Cameras Nuci Med; 35:7-15, 1994
12. Obi J.C1. and Imainvan A.A "Decision Support System for the Intelligent Identification of Alzheimer's using Neuro Fuzzy Logic", International Journal on Soft Computing (IJSC), Vol.2, No.2, PP 25-38, May 2011
13. Sandhya Joshi.et.al "Classification of neuro degenerative Disorders Based on Major Risk Factors Employing Machine Learning Techniques" IACSIT International Journal of Engineering and Technology, Vol.2, No.4, PP 350-355, August 2010
14. Javier Escudero "Machine Learning-Based Method for Personalized and Cost-Effective Detection of Alzheimer's Disease" IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, VOL. 60, NO. 1, PP 164-168, JANUARY 2013
15. Roman Filipovych and Christos Davatzikos "semi supervised classification on medical images-Application to MCI" neuro image Dec, 2012
16. https://www.oasis-brains.org/,https://www.kaggle.com/jboysen/mri-and-Alzheimer's