# DIAGNOSIS AND PREDICT THE STAGES OF ALZHEIMER'S DISEASE BY USING DATA MINING TECHNIQUES

**K. Ambika , C. Gokulavani , S. Harini, S. Jamuna, T. Kiruthika**
Department of Computer Science and Engineering, AVS Engineering College, Salem, Tamilnadu, India.

**ABSTRACT - Alzheimer's is a mind arranged dynamic malady, the reason for this illness called dementia. It makes the synapses of human to be degenerative and pass on. There are various stages called, Mild psychological disability, Mild dementia, Moderate dementia, severe dementia. Neurodegenerative is a significant reason for dementia. An early manifestation of this ailment was not effectively recollecting the ongoing exercises. The people groups with the age over 65 are for the most part influenced. They feel too hard to even think about remembering the ongoing discussions, changes in their conduct, dynamic memory misfortune and changes in their mindset and character, and so on. There are no early manifestations of this ailment. To distinguish these progressions by the doctor by making the two strides called, SAGE and ELISA. In the propelled stage, an individual's misfortune in their cerebrum capacities, similar to parchedness, hunger, and contamination, drives the patients to death. To beat this issue by applying the data mining procedure called K-implies, which is a sort of grouping. It breaks down the patients from the beginning period of the illness. Taking their MRI (Magnetic reverberation imaging) and CT(Computed Tomography) sweeps of a human, shows the loss of cerebrum work which is related to the mass of Alzheimer's. Furthermore, the PET (Positron Emission Tomography) shows the various kinds of cerebrum exercises dependent on the Bloodstream, oxygen level, and glucose level. These outputs are used to figure the decay level in the patients and approach to determine the illness to have exactness.**

**KEYWORDS - SAGE-Self Administered Gerocognitive Examination, ELISA-Enzyme-linked Immune Sorbent Assay., MRI-Magnetic Reverberation Imaging, CT-Computed Tomography, PET-Positron Emission Tomography**

## INTRODUCTION:

Alzheimer's ailment is an irreversible, powerful cerebrum issue that slowly demolishes memory and thinking aptitudes and, over the long haul, the ability to do the most direct endeavors. In a considerable number of individuals with the illness—those with the late-starting sort—indications at first appear in their mid-60s. Simulated intelligence is the investigation of getting PCs to learn and act like individuals do, and improve their learning after some time in the self-overseeing plan, by empowering them with data and information as observations and authentic associations. It epitomizes the ideal objective or extraordinary purpose of AI, as conveyed by various researchers in the field. There is a wide extent of kinds of AI figurings, with hundreds passed on every day, and they're typically assembled by either learning style (for example coordinated learning, autonomous learning, semi-oversaw learning) or by comparability in structure or breaking point (concerning model solicitation, descend into sin, choice tree, gathering, noteworthy learning, and so on.). Despite learning style or cutoff, all mixes of AI figurings include the going with:

- Representation - a lot of classifiers or the language that a PC gets it
- Evaluation - also called objective/scoring limit
- Optimization - search strategy; as often as possible the most imperative scoring classifier, for example; there are both ready to move and custom smoothing out procedures used PC based insight assessment is set up to utilize an arranging illuminating document to make a model. Right when new information knows about the ML assessment, it makes a longing subject to the model. The guess is assessed for accuracy and if the precision is worthy, the Machine Learning calculation is passed on. On the off chance that the precision isn't acceptable, the Machine Learning figuring is set up over and over with an all-inclusive arranging informational record.

**Types:**

1. Supervised
2. Unsupervised
3. Reinforcement Learning

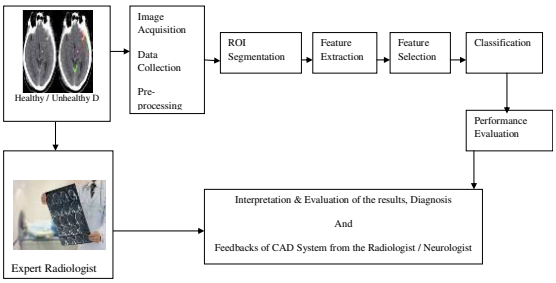
Managed: Regulated Learning is, the spot you can consider the learning is guided by an instructor. We have a dataset that goes about as an educator and its primary duty is to set up the model or the machine. Exactly when the model gets prepared, it can begin picking a craving or choice when new information is given to it.

1. Regression
2. Decision Tree
3. Random Forest
4. Classification

Unaided: It shows the structure of given information exactly when the model is given a dataset, it along these lines discovers models and relationship in the dataset by making packs in it. What it can't do is add engravings to the social occasion.

1. Clustering
2. Association Analysis
3. Hidden Markov Model
4. K-implies
5. Fuzzy
6. Fuzzy KNN
7. ANFIS

To plan through applying an AI to anticipate coronary illness, for example, AD powerfully. AD sickness delineates the degree of conditions that sway your heart. Ailments under the coronary infirmity umbrella wire vein sicknesses, for example, coronary course pain, heart musicality issues (arrhythmias), and heart spurn you're carried into the world with (regular heart deserts), among others. The cardiovascular malady, for the most part, suggests conditions that join limited or hindered veins that can impel a coronary frustration, chest torture (angina), or stroke. Other heart conditions, for instance, those that impact your heart's muscle, valves, or musicality, similarly are seen as sorts of coronary illness.it is difficult to perceive coronary ailment because of a couple of contributory peril factors, for instance, diabetes, hypertension, raised cholesterol, abnormal heartbeat rate, and various diverse elements.AI (ML) shows to be feasible in assisting with making decisions and gauges, it is used to anticipate the individualls from the coronary ailment and easy to screen and check their prosperity conditions wherever.




***Overall architecture***

**PROPOSED WORK**

**1.K-means:**

**Kmeans** algorithm is an iterative algorithm that tries to partition the dataset into *K*pre-defined distinct non-overlapping subgroups (clusters) where each data point belongs to **only one group**. It tries to make the inter-cluster data points as similar as possible while also keeping the clusters as different (far) as possible. It assigns data points to a cluster such that the sum of the squared distance between the data points and the cluster's centroid (arithmetic mean of all the data points that belong to that cluster) is at the minimum. The less variation we have within clusters, the more homogeneous (similar) the data points are within the same cluster.

The way the K-means algorithm works is as follows:

1. Specify the number of clusters *K*.
2. Initialize centroids by first shuffling the dataset and then randomly selecting *K* data points for the centroids without replacement.
3. Keep iterating until there is no change to the centroids. i.e assignment of data points to clusters isn't changing.

- Compute the sum of the squared distance between data points and all centroids.

- Assign each data point to the closest cluster (centroid).

- Compute the centroids for the clusters by taking the average of all data points that belong to each cluster.

The objective function is:

$$J = \sum_{i=1}^{m} \sum_{k=1}^{K} w_{ik} \|x^i - \mu_k\|^2 \quad (1)$$

where wik=1 for data point xi if it belongs to cluster *k*; otherwise, wik=0. Also, μk is the centroid of xi's cluster.

It's a minimization problem of two parts. We first minimize J w.r.t. wik and treat μk fixed. Then we minimize J w.r.t. μk and treat wik fixed. Technically speaking, we differentiate J w.r.t. wik first and update cluster assignments (*E-step*). Then we differentiate J w.r.t. μk and recompute the centroids after the cluster assignments from the previous step (*M-step*). Therefore, E-step is:

$$\frac{\partial J}{\partial w_{ik}} = \sum_{i=1}^{m} \sum_{k=1}^{K} \|x^i - \mu_k\|^2$$

$$\Rightarrow w_{ik} = \begin{cases} 1 & \text{if } k = argmin_j \|x^i - \mu_j\|^2 \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

**Few things to note here:**

- Since clustering algorithms including K-means use distance-based measurements to determine the similarity between data points, it's recommended to standardize the data to have a mean of zero and a standard deviation of one since almost always the features in any dataset would have different units of measurements such as age vs income.

- Given kmeans iterative nature and the random initialization of centroids at the start of the algorithm, different initializations may lead to different clusters since the K-means algorithm may be *stuck in a local optimum and may not converge to the global optimum*. Therefore, it's recommended to run the algorithm using different initializations of centroids and pick the results of the run that that yielded the lower sum of squared distance.

The approach kmeans follow to solve the problem is called **Expectation-Maximization**. The E-step is assigning the data points to the closest cluster. The M-step is computing the centroid of each cluster. Below is a break down of how we can solve it mathematically (feel free to skip it).

In other words, assign the data point xi to the closest cluster judged by its sum of squared distance from the cluster's centroid.

And M-step is:

$$\frac{\partial J}{\partial \mu_k} = 2 \sum_{i=1}^{m} w_{ik} \left(x^i - \mu_k\right) = 0$$

$$\Rightarrow \mu_k = \frac{\sum_{i=1}^{m} w_{ik} x^i}{\sum_{i=1}^{m} w_{ik}} \quad (3)$$

Which translates to recomputing the centroid of each cluster to reflect the new assignments.

- Assignment of examples isn't changing is the same thing as no change in within-cluster variation:

$$\frac{1}{m} \sum_{i=1}^{m} \|x_i - \mu_{c_i}\|_2 \quad (4)$$

**Applications**

K-means algorithm is very popular and used in a variety of applications such as market segmentation, document clustering, image segmentation and image compression, etc. The goal usually when we undergo a cluster analysis is either

Get a meaningful intuition of the structure of the data we're dealing with.

1. Cluster-then-predict where different models will be built for different subgroups if we believe there is a wide variation in the behaviors of different subgroups. An example of that is clustering patients into different subgroups and build a model for each subgroup to predict the probability of the risk of having a heart attack.

In this post, we'll apply clustering in two cases:

- Geyser eruptions segmentation (2D dataset).
- Image compression.

**K means on Geyser's Eruptions Segmentation**

We'll first implement the K-means algorithm on the 2D dataset and see how it works. The dataset has 272 observations and 2 features. The data covers the waiting time between eruptions and the duration of the eruption for the Old Faithful geyser in Yellowstone National Park, Wyoming, USA. We will try to find *K* subgroups within the data points and group them accordingly. Below is the description of the features:

- eruptions (float): Eruption time in minutes.
- waiting (int): Waiting time to next eruption.

**1. FUZZY:**

**The fuzzy logic algorithm** helps to solve a problem after considering all available data. Then it takes the best possible decision for the given input. The FL method imitates the way of decision making in a human which considers all the possibilities between digital values T and F.

The term fuzzy means things that are not very clear or vague. In real life, we may come across a situation where we can't decide whether the statement is true or false. At that time, fuzzy logic offers very valuable flexibility for reasoning. We can also consider the uncertainties of any situation.

However, fuzzy logic is never a cure for all. Therefore, it is equally important to understand where we should not use fuzzy logic.

Here, are certain situations when you better not use Fuzzy Logic:

- If you don't find it convenient to map an input space to an output space
- Fuzzy logic should not be used when you can use common sense
- Many controllers can do a fine job without the use of fuzzy logic

**Characteristics of Fuzzy Logic**

Here, are some important characteristics of fuzzy logic:

- Flexible and easy to implement machine learning technique

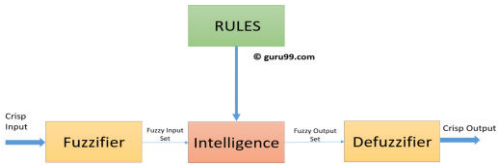
Helps you to mimic the logic of human thought

- Logic may have two values which represent two possible solutions
- A highly suitable method for uncertain or approximate reasoning
- Fuzzy logic views inference as a process of propagating elastic constraints
- Fuzzy logic allows you to build nonlinear functions of arbitrary complexity.

- Fuzzy logic should be built with the complete guidance of experts

**Rule-Based:**It contains all the rules and the if-then conditions offered by the experts to control the decision-making system. The recent update in the fuzzy theory provides various methods for the design and tuning of fuzzy theory provides various methods for the design and tuning of fuzzy controllers. These updates significantly reduce the number of fuzzy sets of rules.

**Fuzzy Logic Architecture**




Fuzzy Logic architecture has four main parts as shown in the diagram:

theory provides various methods for the design and tuning of fuzzy controllers. These updates significantly reduce the number of fuzzy sets of rules.

**Fuzzification:**

The fuzzification step helps to convert inputs. It allows you to convert, crisp numbers into fuzzy sets. Crisp inputs measured by sensors and passed into the control system for further processing. Like Room temperature, pressure, etc.

**Inference Engine:**

It helps you to determines the degree of match between fuzzy input and the rules. Based on the % match, it determines which rules need implement according to the given input field. After this, the applied rules are combined to develop the control actions.

**Defuzzification:**

At last, the Defuzzification process is performed to convert the fuzzy sets into a crisp value. There are many types of techniques available, so you need to select it which is best suited when it is used with an expert system.

**Advantages of Fuzzy Logic System**

- The structure of Fuzzy Logic Systems is easy and understandable
- Fuzzy logic is widely used for commercial and practical purposes
- It helps you to control machines and consumer products
- It may not offer accurate reasoning, but the only acceptable reasoning
- It helps you to deal with the uncertainty in engineering
- Mostly robust as no precise inputs required
- It can be programmed to in the situation when the feedback sensor stops working
- It can easily be modified to improve or alter system performance
- inexpensive sensors can be used which helps you to keep the overall system cost and complexity low
- It provides the most effective solution to complex issues

**Disadvantages of Fuzzy Logic Systems**

- Fuzzy logic is not always accurate, so the results are perceived based on assumptions, so it may not be widely accepted.

- Fuzzy systems don't have the capability of machine learning as-well-as neural network type pattern recognition

**3.FUZZY KNN**

A **fuzzy K-nearest neighbourr algorithm**. ... The theory of **fuzzy** sets is introduced into the **K -nearest neighbor** technique to develop a **fuzzy** version of the **algorithm**. Three methods of assigning **fuzzy** memberships to the labeled samples are proposed, and experimental results and comparisons to the crisp version are presented.

The fuzzy K-nearest algorithm, as far as I can tell, is a fuzzy implementation of kNN. The differential between the two is fuzzy set membership allocation. This means, that, instead of clear set designation, in terms of disposition and Bayesian prediction. Now, implementations of this - I could not find. But I am guessing, that the inherent implementation is something akin to a semi-stochastic differential in terms of Sigmatic ranges or something akin. It'd be in line with every other algorithm I have seen in the field, at least.

The **k-nearest neighbors** (**KNN**) **algorithm** is a simple, easy-to-implement supervised **machine learning algorithm** that can be used to solve both classification and regression problems. K-nearest neighbors (KNN) algorithm is a type of supervised ML algorithm that can be used for both classifications as well as regression predictive problems. However, it is mainly used for classification predictive problems in the industry. The following two properties would define KNN.

- **Lazy learning algorithm** − KNN is a lazy learning algorithm because it does not have a specialized training phase and uses all the data for training while classification.

- **Non-parametric learning algorithm** − KNN is also a non-parametric learning algorithm because it doesn't assume anything about the underlying data.

**Working of KNN Algorithm**

K-nearest neighbors (KNN) algorithm uses 'feature similarity' to predict the values of new datapoints which further means that the new data point will be assigned a value based on how closely it matches the points in the training set. We can understand its working with the help of following steps −

**Step 1** − For implementing any algorithm, we need a dataset. So during the first step of KNN, we must load the training as well as test data.

**Step 2** − Next, we need to choose the value of K i.e. the nearest data points. K can be any integer.

**Step 3** − For each point in the test data do the following −

- **3.1** − Calculate the distance between test data and each row of training data with the help of any of the method namely: Euclidean, Manhattan or Hamming distance. The most commonly used method to calculate distance is Euclidean.

- **3.2** − Now, based on the distance value, sort them in ascending order.

- **3.3** − Next, it will choose the top K rows from the sorted array.

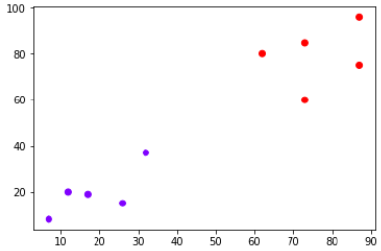
- **3.4** − Now, it will assign a class to the test point based on the most frequent class of these rows.

**Step 4** − End

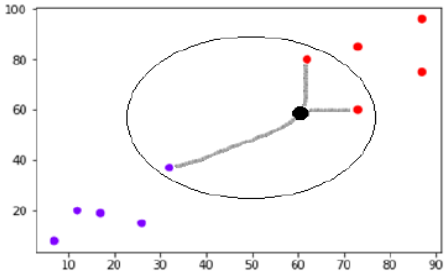
**Example**

The following is an example to understand the concept of K and working of KNN algorithm −

Suppose we have a dataset which can be plotted as follows −



Now, we need to classify a new data point with a black dot (at point 60,60) into a blue or red class. We are assuming K = 3



We can see in the above diagram the three nearest neighbors of the data point with a black dot. Among those three, two of them lie in Red class hence the black dot will also be assigned in red class.

**2.ANFIS**

An adaptive neuro-fuzzy inference system or adaptive network-based fuzzy inference system (**ANFIS**) is a kind of artificial neural network that is based on Takagi–Sugeno fuzzy inference system. The technique was developed in the early 1990s. ... Hence, **ANFIS** is considered to be a universal estimator.

Fuzzy Inference System is the key unit of a fuzzy logic system having decision making as its primary work. It uses the "IF…THEN" rules along with connectors "OR" or "AND" for drawing essential decision rules.

**Characteristics of Fuzzy Inference System**

Following are some characteristics of FIS −

- The output from FIS is always a fuzzy set irrespective of its input which can be fuzzy or crisp.

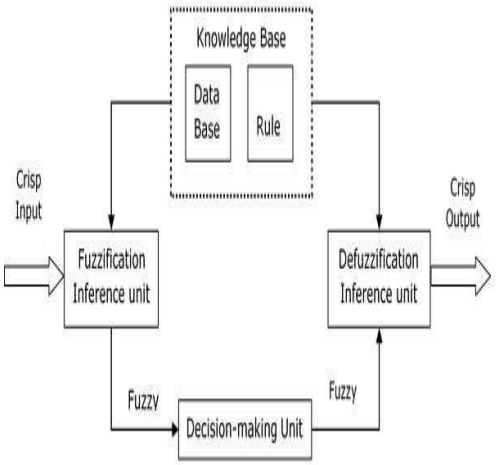
i.e. it would find three nearest data points. It is shown in the next diagram −

- It is necessary to have a fuzzy output when it is used as a controller.
- A defuzzification unit would be there with FIS to convert fuzzy variables into crisp variables.

**Functional Blocks of FIS**

The following five functional blocks will help you understand the construction of FIS −

- **Rule Base** − It contains fuzzy IF-THEN rules.
- **Database** − It defines the membership functions of fuzzy sets used in fuzzy rules.
- **Decision-making Unit** − It performs an operation on rules.
- **Fuzzification Interface Unit** − It converts the crisp quantities into fuzzy quantities.
- **Defuzzification Interface Unit** − It converts the fuzzy quantities into crisp quantities. Following is a block diagram of a fuzzy interference system.




**Working of FIS**

The working of the FIS consists of the following steps −

- A fuzzification unit supports the application of numerous fuzzification methods and converts the crisp input into fuzzy input.
- A knowledgebase - a collection of rule base and database is formed upon the conversion of crisp input into fuzzy input.
- The defuzzification unit fuzzy input is finally converted into crisp output.

**Methods of FIS**

Let us now discuss the different methods of FIS. Following are the two important methods of FIS, having different consequent of fuzzy rules −

- Mamdani Fuzzy Inference System
- Takagi-Sugeno Fuzzy Model (TS Method)

**Mamdani Fuzzy Inference System**

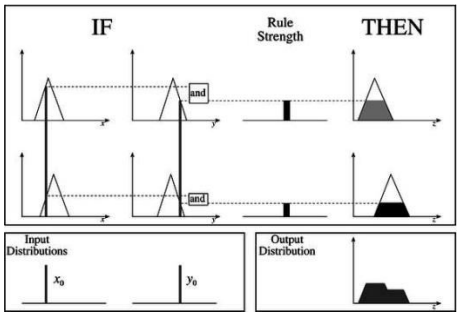
This system was proposed in 1975 by EbhasimMamdani. Basically, it was anticipated to control a steam engine and boiler combination by synthesizing a set of fuzzy rules obtained from people working on the system.

**Steps for Computing the Output**

Following steps need to be followed to compute the output from this FIS −

- **Step 1** − Set of fuzzy rules that need to be determined in this step.
- **Step 2** − In this step, by using the input membership function, the input would be made fuzzy.
- **Step 3** − Now establish the rule strength by combining the fuzzified inputs according to fuzzy rules.
- **Step 4** − In this step, determine the consequences of rule by combining the rule strength and the output membership function.
- **Step 5** − Forgetting output distribution combines all the consequents.
- **Step 6** − Finally, a defuzzified output distribution is obtained.

Following is a block diagram of the Mamdani Fuzzy Interface System.




**Takagi-Sugeno Fuzzy Model (TS Method)**

This model was proposed by Takagi, Sugeno and Kang in 1985. Format of this rule is given as −

*IF x is A and y is B THEN Z = f(x,y)*

Here, *AB* is fuzzy sets in antecedents and $z = f(x,y)$ is a crisp function in the consequent.

**Fuzzy Inference Process**

The fuzzy inference process under Takagi-Sugeno Fuzzy Model (TS Method) works in the following way −

- **Step 1: Fuzzifying the inputs** − Here, the inputs of the system are made fuzzy.

- **Step 2: Applying the fuzzy operator** − In this step, the fuzzy operators must be applied to get the output.

**Rule Format of the Sugeno Form**

The rule format of Sugeno form is given by −

*if 7 = x and 9 = y then output is z = ax+by+c*

**Comparison between the two methods**

Let us now understand the comparison between the Mamdani System and the Sugeno Model.

- **Output Membership Function** − The main difference between them is on the basis of the output membership function. The Sugeno output membership functions are either linear or constant.

- **Aggregation and Defuzzification Procedure** − The difference between them also lies in the consequence of fuzzy rules and due to the same their aggregation and defuzzification procedure also differ.

- **Mathematical Rules** − More mathematical rules exist for the Sugeno rule than the Mamdani rule.

**Adjustable Parameters** − The Sugeno controller has more adjustable parameters than the Mamdani controller.

**RESULT AND DISCUSSION:**

By using FUZZY, FUZZY KNN it will be producing a result for the 4 possibilities which is True positive, True negative, False positive, False negative, Then we have to identify means crowl revely which the help of the classify algorithm then feature selection and feature extraction will be apply for the answers and clonolg it have been producing result accuracy 99% successfully. The result experimental evaluation.

**B. TABLE STRUCTURE**

**TABLE DESIGN**

**Table 1:** Main Form

**Table 2: AM**

| Field Name | Data Type | Size |
|---|---|---|
| K-Means | varchar | 11 |
| SD, MSE | varchar | 30 |

| Field Name | Data Type | Size |
|---|---|---|
| S. NO | varchar | 11 |
| Name | varchar | 250 |

**Table 3: CONVERSION**

| Field Name | Data Type | Size |
|---|---|---|
| Stages | Varchar | 30 |
| AD Detection | Varchar | 30 |

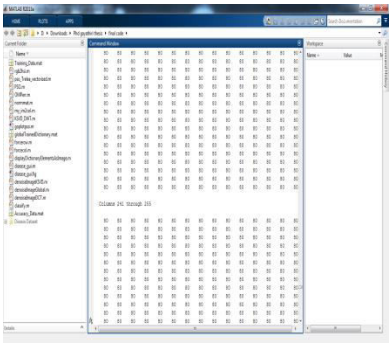
**Table 4: NodeMCU**

| Field Name | Data Type | Size |
|---|---|---|
| True | Int | 255 |
| False | varchar | 1000 |

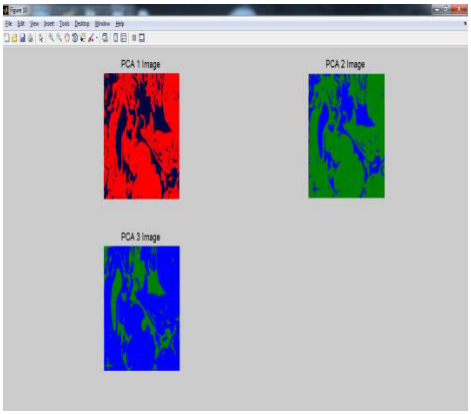
**Table 5: Result**

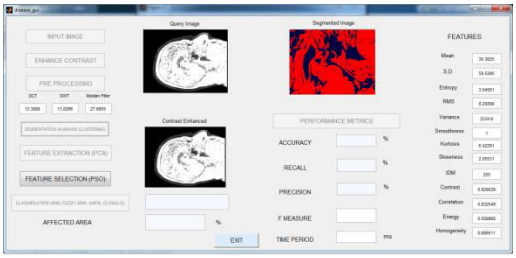
| Field Name | Data Type | Size |
|---|---|---|
| ANFIS | varchar | 30 |
| AD Detection | varchar | 255 |

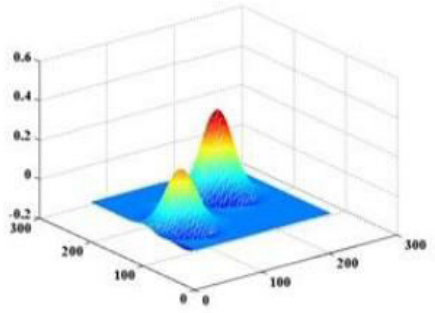
**SAMPLE INPUT**

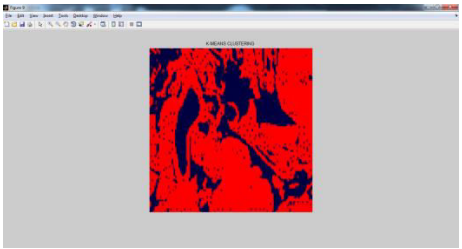


**SAMPLE OUTPUT**





## CONCLUSION AND FUTURE ENHANCEMENT

Three objectives of data mining were defined based on both the exploration of Alzheimer's disease dataset and the objectives of this research. They were evaluated against the trained models. Results showed that all three models had achieved the stated objectives, suggesting that they could be used to provide decision support to medical doctors for diagnosing patients and discovering medical factors associated with Alzheimer‟s Disease. The objectives were as follows: The first objective was to discover the significant influences and relationships in the medical inputs associated with the predictable state of Alzheimer's disease. The Dependency viewer in Association Rules, Decision Trees, and Naïve Bayes models showed the results from the most significant to the least significant medical predictors. Medical Doctors can use this information to further analyze the strengths and weaknesses of the medical attributes associated with Alzheimer's disease. The second objective was to predict those who are likely to be diagnosed with Alzheimer's disease, given patients‟ medical profiles. It was found that all models were able to perform this task using a singleton query, which made use of single input cases and multiple input cases respectively and also to show the rate of the disease. The third objective: was to compare the different data mining techniques. It was found that the Decision Tree model was the best in the diagnosis process and the Naïve Bayes model was the best in identifying the characteristics of patients with Alzheimer's disease and showing the probability of each input attribute for the predictable state.

Based on the previous conclusions and several issues that arose during the study, some topics may be considered as future opportunities to be explored by interested researchers. They are the following:

[1] Using additional data mining techniques, such as Genetic Algorithms in the predicting phase of Alzheimer's disease.

[2]Adding patient data related to MRI and/or CT scan to get more accurate results.

[3] Using the output of the model in the development of an expert system for the diagnosis and prognosis of the disease.

## REFERENCES:


[1] Alzheimer’s Association, “2017 Alzheimer’s Disease Facts and Figures,” 2017.

[2] C. Reitz and R. Mayeux, “Alzheimer disease: Epidemiology, diagnostic criteria, risk factors and biomarkers,” Biochemical Pharmacology, vol. 88, no. 4. pp. 640–651, 2014.

[3] M. DiLuca and J. Olesen, “The cost of brain diseases: A burden or a challenge?,” Neuron, vol. 82, no. 6. pp. 1205–1208, 2014.

[4] A. Wimo, L. Jönsson, J. Bond, M. Prince, and B. Winblad, “The worldwide economic impact of dementia 2010,” Alzheimer’s Dement., vol. 9, no. 1, pp. 1–11, 2013.

[5] G. Vradenburg, “A pivotal moment in Alzheimer’s disease and dementia: How global unity of purpose and action can beat the disease by 2025,” Expert Review of Neurotherapeutics, vol. 15, no. 1. pp. 73–82, 2014.

[6] K. Iqbal and I. Grundke-Iqbal, “Alzheimer’s disease, a multifactorial disorder seeking multitherapies,” Alzheimer’s and Dementia, vol. 6, no. 5. pp. 420–424, 2010.

[7] L. S. Schneider et al., “Clinical trials and late stage drug development for Alzheimer’s disease: An appraisal from 1984 to 2014,” J. Intern. Med., vol. 275, no. 3, pp. 251–283, 2014.

[8] J. Cummings, G. Lee, T. Mortsdorf, A. Ritter, and K. Zhong, “Alzheimer’s disease drug development pipeline: 2017,” Alzheimers Dement (N Y), vol. 3, no. 3, pp. 367–384, 2017.

[9] H. M. Snyder et al., “Developing novel blood-based biomarkers for Alzheimer’s disease,” Alzheimer’s and Dementia, vol. 10, no. 1. pp. 109–114, 2014.

[10] J. Godyn, J. Jonczyk, D. Panek, and B. Malawska, “Therapeutic strategies for Alzheimer’s disease in clinical trials,” Pharmacol. Reports, vol. 68, no. 1, pp. 127–138, 2016