

# Machine Learning based Chat Bot Recommendations for Mental Health Diseases

S. Senthil, PhD

School of Computer Science and  
Application  
REVA University, Bangalore, India

Immanuel D.

School of Computer Science and  
Application  
REVA University, Bangalore, India

Aaquib Nawaz S.

Senior Software Engineer  
ACI Worldwide Pvt Ltd  
Bangalore, India

## ABSTRACT

A chat bot based totally system that may use machine getting to know strategies that allows you to answers the questions of the customers in a computerized fashion for universal questions. examine the chats of the person so as expect the form of medical ailment the consumer is having the use of Naive Bayes method along with natural language processing using schooling key phrases generated with the aid of analysis of top intellectual fitness web sites maintained by using doctors and hospitals. once the prediction of sickness is finished find the level of the ailment with the aid of having signs and symptoms based totally chat among the person and chat bot the use of single dimension k-means algorithm and aid vector machine aggregate. If the person is having low stage or medium degree kind of intellectual ailment the research can provide guidelines from professional docs with different treatment plans for every disease degree type.

## Keywords

Mental health, data mining, machine learning, chat bot

## 1. INTRODUCTION

The recognition and diagnosis for the dementia at an initial stage is done by either based on symptoms by the user or generally first contact friends or relatives. Mild Cognitive Impairment (MCI) is used to find the dementia patient but takes usually few years to find the root cause as

lot of symptoms are ignored thinking it is due to age[2]. Due to delayed discovery the doctors are unable to recommend a definitive kind of treatments and hence write letters for cognitive assessments, expensive neurosurgical [3]. Magnetic Resonance Imaging (MRI) along with Positron emission Tomography (PET) are the recommended scans by various doctors and this results need to be evaluated by experts in the field of radiology and anatomy who will then analyse the scans and perform certain manual tasks which are time consuming in nature and there is a variation for each section of the brain [4]-[5]

The use of neurosurgical applications. The foremost challenge being cost at £169 and £844 per patient within the UK for MRI and PET scans respectively. The scanning tests are very costly and might cost in the range of 169 to 844 euros per person [6]. This cost can be reduced by using alternative and readily available tests. There is a lot of social misinformation which is increasing the refusals rates for future treatment of dementia patients after initial analysis which will cause a proportional increase in other factors like anxiety, social stigma [7].

SVM makes use of training data and then creates hyper planes in order to group the data and then predict the class label by performing model generation along with distance computation

on the hyper planes. In order to have better accuracy the separated distance between planes must be maximum which can be computed using Lagrangian distribution [8].

In a Random Forest (RF) based methods multiple independent decision trees are created and then set of features are used as attributes which will calculate independent decision trees and then predict the class. The class labels from each of the decision trees are taken out and then final prediction is done based on average or manual voting [9]

## PROMBLEM STATEMENT

In today's world the number of patients is increasing day by day and then the patients are waiting in the queue to meet doctors and sometimes they get appointment and few times they don't get appointment. More over the patients who are depressed need to meet the doctor at regular intervals. Hence a system is needed in which the patient can get registered and then perform the chatting based on the conversations the bot system can provide the suggestions then also the chat bot system can even provide suggestions at the end of chatting

## 2. BACKGROUND

The recognition and diagnosis for the dementia at an initial stage is done by either based on symptoms by the user or generally first contact friends or relatives. Mild Cognitive Impairment (MCI) is used to find the dementia patient but takes usually few years to find the root cause as lot of symptoms are ignored thinking it is due to age [2]

Due to delayed discovery the doctors are unable to recommend a definitive kind of treatments and hence write letters for cognitive assessments, expensive neurosurgical [3]

Chat bot based systems are very helpful for making a connection between the community and chat based system. A series of known questions are feeder into the bot for training purposes and then a match is performed based on bigram between the chat done by the user and the trained sentences. If the match is exact or meets certain threshold then an automated suggestion is replied to the end user [10].

Domain based automated answer systems will be trained with certain questions for each domain. The drawback of this

system is that there should be an exact match of questions with the trained question and if there is match then reply is sent. Otherwise ontology based execution is done in order to derive the partially matched question along with reply [11].

## 3. PROPOSED METHOD

In the proposed approach an artificial intelligence based chat bot is used to assist the patients in which the patient will ask a question, the automated bot will ask series of questions related to a disease, based on the answers the analysis is performed,

classification of the disease is performed for the end user. If the disease level is highest then appointment is given otherwise only suggestions are given to the end user. The classification graphs are obtained based on gender and age range. The scope of the work is to build an end to end web application which allows the users to register, allows the admin to create more admin or doctors. The application must have capability to give chat system for the user to get the series of automated question and then based on the answer the analysis is performed on those answers. Once the analysis is completed then the user will be able to get suggestions given by the doctor and then if the condition is critical then user will be able to get the list of suggestions and an appointment request. The approval cycle of the appointments and removal of appointments is also in the scope.

Fig hints at the structure of the machine. As proven in the fig there are numerous components which are concerned inside the working model for the assignment.

### **Angular/ext js view**

This module is billable for front perspectives technology quit using angular and ext js framework alongside java server pages.

### **Tomcat web container**

There are many servers available on the market, which is in charge of handling Web queries. Most contrary servers are cumbersome and are also of a commercial nature. Here we make use of the open source and light server tomcat.

### **Middle Ware – Controlling layer**

This module is responsible for the management of the Internet query and its routing to the authentication layer. It also affects fundamental validations such as empty appraisals and regex validations. If any validation fails, then the response is shipped to the front cease otherwise the request is forwarded to the authentication layer and respective services

### **Artificial Intelligence engine**

The artificial intelligence engine is used to get trained for a set of sentences related to various diseases and send back the responses and also the artificial intelligence engine is trained with specific set of questions and answers which the engine will send depending on the kind of disease the user association happens.

### **Data layer**

The data layer is answerable for garage of records related to registered users, admin, doctors. The statistics layer can also be able to store solutions of customers, appointment statistics in addition to the class records..

### **Authentication layer**

This layer is responsible for viewing the customer's application verification if the application has valid ID for the application and a valid session. The user may be dismissed if the consultation is not valid.

### **Registration service**

This module is responsible for allowing taking the parameters like username, password, first name, last name, email, age and gender, performs the mandatory validations, length validations, regex validations, similar user id and email id validations.

### **Login service**

This module is responsible for taking username and password as an input, perform basic and regex validations. Once the validations are successful then perform the validation of authentication credentials. Once the authentication of the user is successful then login will be successful.

### **Appointment service**

This service is responsible for performing the appointment workflow for all the three kinds of actors namely doctor, admin and patient. Once the user is categorized as having highest level of disease. Then appointment request is generated. The appointment workflow goes to admin, undergoes the approval cycle and users will get the notifications. All these workflow is maintained by the appointment service.

### **Chat bot service**

This chat bot service is responsible for answering the questions asked by the user by making use of the training data. If any of the comment by the user does not match the end user trained query then generic message is given to the end user.

This module is responsible for training the system with the set of questions and answers of each question are provided. If the chat asked by the user does not match then a TF-IDF algorithm is computed to find the similarity.

Responsible for maintaining the set of answers of the chat session.

Responsible for asking series of questions to the end user based on each of the diseases.

### **History profile service**

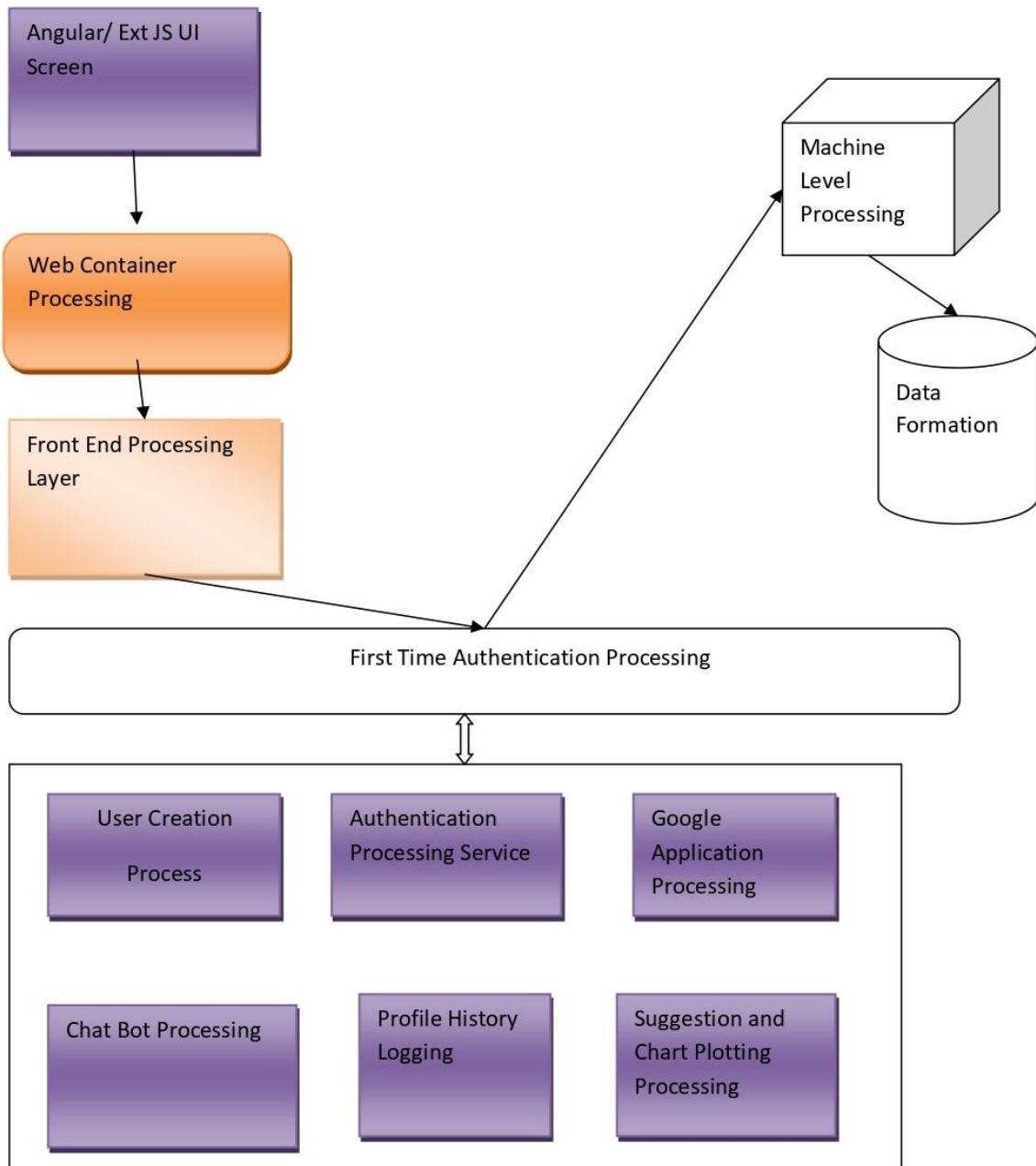
This module is responsible for generating the history profile. The history profile is responsible for maintaining the number of chat bot sessions of the x axis and then disease levels on the y axis and several such disease levels will be generated.

### **Classification, Suggestion and Graph Service**

This module is responsible for performing the computation of weightage for the disease attribute levels and then classifies the disease level for the user. It also is responsible for generating the classification graph data for classification by age and gender

## **4. DISEASE CLASSIFICATION AND LABEL ASSIGNMENT**

This section describes the processing algorithms used in converting the chat into the stream of sequenced sentences. Find the count of sentences and then sentence is divided into words, find the generic question if available in the library then automated answer is provided. Naive Bayes method will be trained by using keywords belonging to each of the category. The class label will be determined by computing mean, standard deviation, probability of each categorizes are found and then which category the probability is maximum is the determined class. Once the disease category has been found the level is found using weighted average method.



**Fig 1: Steps involved in the Proposed Method**

**Weighted Average Measure Algorithm**

The weighted average measure algorithm has the following steps

List of questions under each symptom type will be created and each of the questions will have the following

QuesID- An Autogenerated ID for the Question

QUESDESC- This is the question description

ANSWER1- This is the Answer1 visible to User

ANSWER2- This the Answer2 visible to User

ANSWER3- This the Answer3 visible to User

ANSWER4- This the Answer4 visible to User

DISEASETYPE- This is the type of the disease which

**Table 1: Symptom Answer Trained System**

QuestId	Answer1	Answer2	Answer3	Answer4	Rating1	Rating2	Rating3	Rating4	Disease Type	Selected Answer
1	A1	A2	A3	A4	3	2	1	4	Dementia	A4
2	A1	A2	A3	A4	4	3	2	1	Dementia	A2
3	A1	A2	A3	A4	3	4	1	2	Dementia	A3

The answers given by the user will have the following matrix  
Based on the answers given by the user the total rating will be computed

For example- for the user above the following is the total score is obtained which is given by the following

$$TR = A4 \text{ rating} + A2 \text{ rating} + A3 \text{ rating} = 4 + 3 + 1 = 8$$

The max rating possible is computed which in this case is  $3 * 4 = 12$  i.e L maximum value

The entire range is divided into the following

$$\text{Threshold1} = L / 3;$$

$$\text{Threshold2} = L / 2;$$

$$\text{Threshold3} = L;$$

If the total rating is between  $0 - L/3 - 1$  the user will be treated to have Level1

If the total rating is between  $L/3$  to  $L/2 - 1$  the user will be treated to have Level2

If the total rating is above  $L/2$  the user will be having Level3

Based on the Level of the user the suggestions are extracted and given to the end user.

## 5. RESULT

The entire application is designed using spring based framework and for the front end pages dynamic java server pages along with angular or ext js application. The deployment of web application is done on the web container such as apache tomcat and back end processing is done with the help of MYSQL

Figure 3 shows the welcome page for application in which there are two panels with first panel used for accessing the application and then second panel is used for allowing the registration of users can be done .The accessing process for the application is shown in the Figure 4 in which personal details along with unique access parameter and email will be used can be anxiety, depression and dementia

RATING1- This is the rating for answer1

RATING2- This is the rating for answer2

RATING2- This is the rating for answer3

RATING2- This is the rating for answer4

The users will be asked a list of questions and based on the answers the analysis is performed

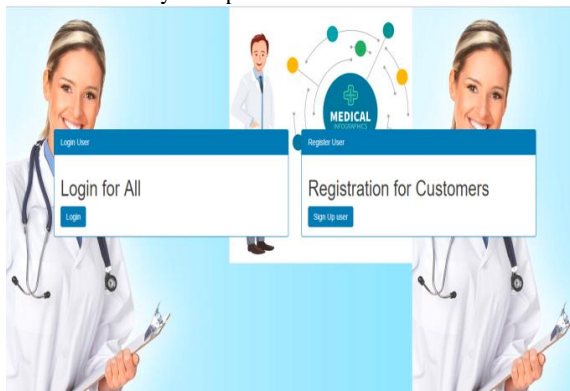


Fig 2: Welcome page for application

Fig 3: User System Access Register Flow

Figure 3 shows the user system access by entering the personal details along with demographic details. From the Figure 3 the patient is registering into the application by providing user name, phone number, gender, email and age which are personal details and then demographic details like city, state and country are provided to perform the registration. If the username “patient12” does not exists then patient will be allowed to register otherwise registration will fail. Also the email id must be unique

## Suggestion Creation By Doctor

Fig 4: Suggestions Created by Doctor

Figure 4 shows the screen which is used by doctor to create the suggestions for selected disease type and level. The doctor has selected clinical depression as the disease type and this is followed by level of the disease which is Level1

Keyword	Type of Keyword
anxiety	anxiety
anxiety	anxiety
anxious	anxiety
you re	anxiety
feeling	anxiety
worried	anxiety
problem	anxiety
work	anxiety
boss	anxiety
butterflies	anxiety
stomach	anxiety
waiting	anxiety
results	anxiety
medical	anxiety
test	anxiety
nervous	anxiety
driving	anxiety
home	anxiety
rush hour	anxiety
traffic	anxiety
cars	anxiety
speed	anxiety
.....	anxiety

**Fig 5: Training Keywords for Disease Kind**

Figure 5 shows the training words used for making the learning for the algorithm. Like this different category of words are used to train the system.

Create User/Doctor/Patient

<b>User Name:</b>	<b>Password:</b>	<b>Email address:</b>
<input type="text" value="doctor"/>	<input type="password" value="....."/>	<input type="text" value="doctor@gmail.com"/>
<b>Phone No:</b>	<b>Gender:</b>	<b>Country</b>
<input type="text" value="7259268905"/>	<input type="text" value="Male"/>	<input type="text" value="India"/>
<b>State</b>	<b>City</b>	<b>Age</b>
<input type="text" value="Karnataka"/>	<input type="text" value="Bangalore"/>	<input type="text" value="23"/>

**Fig 6: Doctor or Admin Access Allocation**

Figure 6 shows the screen or application which is responsible for creating an admin or a doctor or a patient who visits the clinic. The doctor can access functions like creating suggestions, viewing suggestions, viewing appointments and closing the appointment. Admin access is also created which can perform various functions like creating users, approval of users.

Chat For Health Care

hi

How are you? Are you feeling depressed or anxious or Memory loss

**Send Message:**

**Fig 7: Initial Chat Workflow**

Figure 7 shows the chat performed by the end user. The user sends a ‘Hi’ and automated chat bot sends a generic reply with a value “How are you? Are you feeling depressed or anxiety or memory loss”. This reply is from the generic chat answer from the trained phrases.

Hi

How are you? Are you feeling depressed or anxious or Memory loss

i am feeling depressed

Are you sensitive to freezing  
 a)not at all  
 b)yes- 100%  
 c)yes-50%  
 d)<50%  
 Please Provide answer as either a or b or c or d

**Send Message:**

**Fig 8: Symptom based Question**

Once the user has been classified as belonging to specific disease then based on the disease type a symptom based question is asked for the end user. The user will type the answer in the send message text box and clicks on Send Message button which will send the input to the back end process which will send back remaining symptom questions leaving the answered question.

The next figure shows the series of questions asked for the patient and answers being capture in terms of the options which can be either “a” or “b” or “c” or “d”. Figure 9 also shows the suggestions given by the doctor for that specific disease level and that specific category. As seen from Figure 9 the doctor has recommended a tablet and also longer sleep hours.

Based on the answers which are given by the user then level of the user is highest. If the level is highest then appointment button will be shown are shown to the end user and an appointment with no doctor or date is initially created as shown in Figure 10

Do you like oily food?  
a)not at all  
b)yes- 100%  
c)yes-50%  
d)<50%  
Please Provide answer as either a or b or c or d

b

---

How much ghee do you use for 100 gms of food?  
a)yes 1 gm  
b)yes 5 to 10 gms  
c)> 10 gms and <15 gm  
d)> 15 gms  
Please Provide answer as either a or b or c or d

c

---

How many hours you sleep daily  
a)<5 hrs  
b)>5 hrs and <8 hrs  
c)>8hrs and <12 hrs  
d)>12 hrs  
Please Provide answer as either a or b or c or d

d

These is the suggestions for Level( 1) Have clinical840 tablet 3 times 2) Sleep for 16 hours

**Fig 9: Series of Symptom based questions**

Customer Appointment	
User Name	Status
patientsachin	Your Appointment is awaiting Approval

**Fig 10: Appointment for Highest Level User**

The administrator of the clinic or hospital will use the system to check for doctors availability and allocate a time for appointment with in time and out time and once the appointment is granted then the dashboard of the patient will be updated as below

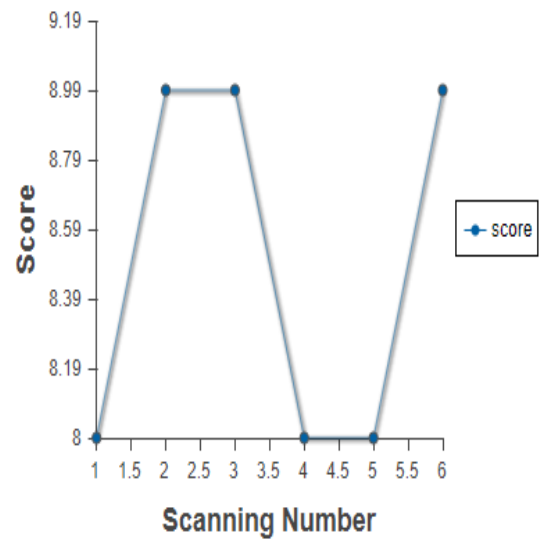
Customer Appointment			
User Name	Status		
patientsachin	Your Appointment has been approved		

Doctor Name	Date	In Time	Out Time
mcadoctor	10/0/2022	11:30AM	12:00PM

**Fig 11: Customer Time Blocked Successfully**

Figure 11 shows that the user has got an appointment for doctor around 11:30 AM on 10<sup>th</sup> of sept 2022.



**Fig 12: Graph of Patient Health Progress**

When ever a test is taken by the patient in the application a unique auto generated scanning number is assigned for each test and then based on symptoms of the diseases the patient is assigned a score. Figure 12 shows the heath progress of the patient. From the graph it is evident that for scanning number 1 the score is 8, scanning number 2 the score is 9, scanning number 3 the score is 9 , scanning number 4 along with 5 it has again reduced to 8 and during the 6<sup>th</sup> iteration the score is 9.

**Table 2 : Time Taken**

Range of Users	Average Time Taken (s)
100-1000	20.587 s
1001- 2000	40.890 s
1002- 5002	200.7654s

Latest Suggestions Updated by Any Doctor		
Disease Type	Level	Suggestion
depression	level1	Get enough sleep. Depression can make it hard to get enough shut-eye, and too little sleep can make depression worse. Have half a sleeping tablet
depression	level2	Eat healthy. There is no magic diet that fixes depression. It's a good idea to watch what you eat, though. If depression tends to make you overeat, getting in control of your eating will help you feel better. Take paracetamol
depression	level3	It temporarily boosts feel-good chemicals called endorphins. It may also have long-term benefits for people with depression. Regular exercise seems to encourage the brain to rewire itself in positive ways When you're depressed, you may want to pull back from life and give up your responsibilities at home and at work. Don't. Staying involved and having daily responsibilities can help you maintain a lifestyle that can help counter depression.

Figure 13: Depression Related Treatment Suggestions by Doctor

anxiety	level1	Exercise is a great way to burn off anxious energy, and research tends to support this use.
anxiety	level2	Meditation can help to slow racing thoughts, making it easier to manage stress and anxiety. A wide range of meditation styles, including mindfulness and meditation Some people unconsciously tense the muscles and clench the jaw in response to anxiety. Progressive relaxation exercises can help. Try lying in a comfortable position and slowly constricting and relaxing each muscle group, beginning with the toes and working up to the shoulders and jaw.
anxiety	level3	Some people feel anxious if they have too many commitments at once. These may involve family, work, and health-related activities. Having a plan in place for the next necessary action can help to keep this anxiety at bay. Effective time management strategies can help people to focus on one task at a time. Book-based planners and online calendars can help, as can resisting the urge to multitask.

Figure 14: Anxiety Related Suggestions provided by Doctor

Table 2 shows the time taken for the range of users. As the user count increases the average time taken for the algorithm execution also increases.

The suggestions provided by doctors for various category and disease level is summarized in Figure 13. Figure 13 shows the different suggestions given by doctor based on the level of the disease for dementia. There are three different levels namely Level1, Level2 and Level3. Level 1 indicates the lowest range

for the disease, Level2 is the mid range and Level3 is the highest range. Hence different treatments are suggested in column 3 for different levels of the disease. Figure 14 is similar to Figure 13 but it corresponds to anxiety related issues. With each level having there own suggestions provided by the doctor

## 6. CONCLUSION

In this project there are 3 types of users namely Admin, User

and Doctor. The user will be able to register into the application and once the registration is successful then the user will be able to do a chat bot and then perform the analysis on the answers in order to generate the category of disease for the user. Once the disease is found out the list of suggestions are generated based on the level for the user. The user will also be allowed to find the list of top clinics and hospitals for the given disease type based on google suggestions. The Admin will be able to create a doctor. The doctor will be able to create suggestions and then view suggestions. The work can be future extended in order determine the kidney stage by asking a list of questions related to other diseases and then generate the health profile of the user

## 7. REFERENCES

- [1] M. Boustani, C. M. Callahan, F. W. Unverzagt, M. G. Austrom, A. J. Perkins, B. A. Fultz, S. L. Hui, and H. C. Hendrie, "Implementing a screening and diagnosis program for dementia in primary care," *Journal of General Internal Medicine*, vol. 20, no. 7, pp. 572–577, 2005.
- [2] Alzheimer's Society, "Dementia UK Report," Tech. Rep., 2018.
- [3] K. Lo´pez-de Ipin˜a, J.-B. Alonso, C. Travieso, J. Sole´-Casals, H. Egiraun, M. Faundez-Zanuy, A. Ezeiza, N. Barroso, M. Eca˜y-Torres, P. Martinez-Lage, and U. Lizardui, "On the Selection of Non-Invasive Methods Based on Speech Analysis Oriented to Automatic Alzheimer Disease Diagnosis," *Sensors*, vol. 13, no. 12, pp. 6730–6745, 2013. Tavel, P. 2007 Modeling and Simulation Design. AK Peters Ltd.
- [4] J. P. Lerch, J. Pruessner, A. P. Zijdenbos, D. L. Collins, S. J. Teipel, H. Hampel, and A. C. Evans, "Automated cortical thickness measurements from MRI can accurately separate Alzheimer's patients from normal elderly controls," *Neurobiology of Aging*, vol. 29, no. 1, pp.23–30, 2008. Forman, G. 2003. An extensive empirical study of feature selection metrics for text classification. *J. Mach. Learn. Res.* 3 (Mar. 2003), 1289-1305.
- [5] S. Kloˆppel, C. M. Stonnington, J. Barnes, F. Chen, C. Chu, C. D. Good, I. Mader, L. A. Mitchell, A. C. Patel, C. C. Roberts, N. C. Fox, C. R. Jack, J. Ashburner, and R. S. J. Frackowiak, "Accuracy of dementia diagnosis - A direct comparison between radiologists and a computerized method," *Brain*, vol. 131, no. 11, pp. 2969–2974, 2008.
- [6] M. Boustani, L. Watson, B. Fultz, A. J. Perkins, and R. Druckenbrod, "Acceptance of dementia screening in continuous care retirement communities: A mailed survey," *International Journal of Geriatric Psychiatry*, vol. 18, no. 9, pp. 780–786, 2003.
- A. So, D. Hooshyar, K. Park, and H. Lim, "Early Diagnosis of De- mentia from Clinical Data by Machine Learning Techniques," *Applied Sciences*, vol. 7, no. 7, p. 651, 2017.
- [7] J. Friedman, T. Hastie, and R. Tibshirani, *The elements of statistical learning. Springer series in statistics* New York, NY, USA:, 2001, vol. 1, no. 10.
- [8] T. N. Tombaugh and N. J. McIntyre, "The mini-mental state examination: a comprehensive review." *Journal of the American Geriatrics Society*, vol. 40, no. 9, pp. 922–935, 1992.
- [9] M. W. Weiner, D. P. Veitch, P. S. Aisen, L. A. Beckett, N. J. Cairns, J. Cedarbaum, M. C. Donohue, R. C. Green, D. Harvey, C. R. Jack, W. Jagust, J. C. Morris, R. C. Petersen, A. J. Saykin, L. Shaw, P. M. Thompson, A. W. Toga, and J. Q. Trojanowski, "Impact of the Alzheimer's Disease Neuroimaging Initiative, 2004 to 2014," *Alzheimer's and Dementia*, vol. 11, no. 7, pp. 865–884, 2015.
- [10] Bayu Setiaji ; Ferry Wahyu Wibowo, "Chatbot Using a Knowledge in Database: Human-to-Machine Conversation Modeling", 2016 7th International Conference on Intelligent Systems, Modelling and Simulation (ISMS), 25-27 Jan. 2016
- [11] Agnese Augello ; Giovanni Pilato ; Alberto Machi ; Salvatore Gaglio, "An Approach to Enhance Chatbot Semantic Power and Maintainability: Experiences within the FRASI Project", 2012 IEEE Sixth International Conference on Semantic Computing, 19-21 Sept. 2012.