



Virtual Therapy using Amazon's Alexa

Augustin Vembil¹, Ayisha Abdul Ravooof², Anagha Merin Sen³

¹Jyothi Engineering College, India, vembilaugustin@gmail.com

²Jyothi Engineering College, India, aayisha005@gmail.com

³Jyothi Engineering College, India, anaghamerin20@gmail.com

ABSTRACT

Today, society has proven us that mental health and Psychological problems are very important and people who are suffering from Depression, Anxiety and PTSD must be helped and shouldn't be left alone. But the sad truth is they are left alone. In Clinical Psychology, such people are prescribed with Counselling and Therapy sessions. There are people who are too reserved to open up to a Psychologist or afraid what society would say if they are found out attending therapy. This project is to build a Virtual Psychologist using Amazon's Alexa. Alexa is an Intelligent Voice Assistant where we can build application in it similar to building android applications for Android OS. The Application is called Therapy AI. It is developed using the Alexa developer console and integrating it with Machine learning Algorithms at different stages of the project. With the help of a real Psychologist we were able to understand the approach for a therapy and the first sessions will be classifying the user's character. For this, we use a simple classifier algorithm on the basis of a questionnaire. Further sessions will be based on this classification. We believe that our project will help a lot of people to open up and cure their depression by using this Application in their personal space and ease people from panic attacks and severe anxiety issues by making them calm.

Key words: Alexa, Sagemaker, Amazon Web Service, Psychotherapy

1. INTRODUCTION

Depression has become a common problem among people of almost all age groups. They step back to open up their problems with someone. In a world of busy and hectic schedules, people do not get time to listen to a person. In such a situation a Virtual Voice Assistant comes in role. Virtual Personal Assistant or AI Assistant is an application program which understands natural language voice commands and does the desired task. There are many such VPAs available and Alexa is one such made by Amazon. Amazon Echo is the first cloud-enabled wireless speaker embedded with a conversational agent that can take voice commands from its users to perform a variety of tasks. In

the context of ubiquitous computing, Echo represents another important advance where technology recedes to the background with a voice interface to interpret and support a user's intentions. Basically, Therapy AI is a skill developed with a voice interface to interpret and support a user's intentions. This is just like developing Android applications in Android OS. There are large number of skills that are already available in their Alexa skill kit. Therapy AI provide the user with therapy sessions as practiced in Clinical Psychology. The first session is a questionnaire that helps in analyzing the user's character and later on further sessions are added accordingly depending on how severe the case is. As the application can be used in personal space, it adds up more convenience to the user by avoiding a visit to a real psychologist also the feeling of what others might think on visiting a psychologist and there is no cost for attending each session.

1.1 ALEXA

Virtual Personal Assistant or AI Assistant is an application program which understands natural language voice commands and does the desired task. There are many such VPAs available and Alexa is one such made by Amazon. Amazon has designed the Alexa Voice Service (AVS) to mimic real conversations and when you ask Alexa any question, you are actually communicating with a cloud-based service. In order to use Alexa, we need a device that integrates the voice technology. Echo, Echo Dot, Echo Show are such devices and this cloud based personal voice assistant has also been integrated into Echo Auto, and with some third-party systems. The commands that Alexa can understand are called Alexa skills, and this seems to grow on a daily basis. Users can create their own skills using Amazon Developer Console and Alexa Skill Kit. There are many software and hardware tools provided by Amazon for its users to learn and create skills. Alexa Skill Kit lets the user to teach Alexa new skills. To develop a new skill for Alexa, user have to login with Amazon account connected to the Alexa device. Then a skill configuration has to be created and this collects the information about the skill such as its name, type of interaction model, the end point or content feed and other information. Alexa service uses this configuration to determine which user request should be

sent to the service for the skill. To build a custom skill, user needs an internet-accessible endpoint for hosting the cloud-based service. For this, it is better to use AWS Lambda which is the web service provided by Amazon. We can write Lambda function in Node.js, Java, Python, C# or Go and web service in any appropriate language. Amazon Simple Storage Service (S3) is the storage provided by Amazon for users to host any images, audio files, video files that the user want to use in their skill.

1.2 HOW AN ALEXA SKILL WORKS?

"ALEXA, OPEN THERAPY AI"

'Alexa' is the wake-up word used to turn on the device. This put Alexa into listening mode and makes it ready to take instructions from users. User can change the wake-up word via Alexa assistant mobile application. 'Open' is the launch word. There are many launch words such as Ask, Search etc. 'Therapy AI' is the invocation name and it is the keyword used to trigger a specific skill. All custom skill needs an invocation name to trigger it. User input contain utterances. Utterances are the phrases the user will use when making a request to Alexa. Alexa identifies the user's intent from the utterance and respond accordingly. The Alexa enabled device sends the user instruction to a cloud-based service called Amazon Voice Service (AVS). AVS is considered as the brain of the Alexa-enabled devices and it performs complex operations such as Natural Language Understanding and Automatic Speech Recognition. AVS processes the response and identify user's intent, then it makes web service request to third party server if needed. AWS Lambda works as the backend and the endpoint linked to the skill takes the request and sends back the response. Echo device translates the response to speech output and delivers to the outside world.

1.3. DEPRESSION

Depression is a state of low mood and aversion to activity. It can affect a person's thoughts, behavior, motivation, feelings, and sense of well-being. It may feature sadness, difficulty in thinking and concentration and a significant increase or decrease in appetite and time spent sleeping. People experiencing depression may have feelings of dejection, hopelessness and, sometimes, suicidal thoughts. It can either be short term or long term. The core symptom of depression is said to be anhedonia, which refers to loss of interest or a loss of feeling of pleasure in certain activities that usually bring joy to people. Depressed mood is a symptom of some mood disorders such as major depressive disorder or dysthymia; it is a normal temporary reaction to life events, such as the loss of a loved one; and it is also a symptom of some physical diseases and a side effect of some drugs and medical treatments.

1.4. PSYCHOTHERAPIC SESSION

The first part of the first session is getting to know one another. If you are able to connect each other, the therapy session would work best. The therapist will need to know why patient is seeking therapy. They may ask what kinds of needs or issues patient would like to address in the treatment together as well as what patient has done to manage his mental health in the past. They want to talk through what worked and what didn't to get an understanding of how best to help him. As a secondary part therapy, therapist may ask some questions. The questions are from Patient health care questionnaire (PHQ-9). The Patient Health Questionnaire (PHQ)-9 is conditionally used to diagnose depression and grade severity of symptoms in general medical and mental health settings.

2. LITERATURE REVIEW

2.1. Alexa-based Real-Time Attendance System

This paper is about building an Alexa based attendance system. The aim is to modernize the conventional attendance system using the AI assistant Alexa. Alexa is easily available and can be used as a personal device at any time. The main goal of this work is to simplify the effort spend by teachers in maintaining the attendance of students. Alexa and Raspberry pi are the main components for this work. (Anush, 2018) Each faculty will have an Alexa enabled device which can be deployed by them to start the process. Initially the faculty have to register their name and passphrase on to the server so that they are validated in to the system. Then, a database is created with a table with the following field; Name, UIN, Total classes, attended classes, percentage and attendance. While the faculty says the initiating command, it will establish the connection to the server hosted on Raspberry pi. This triggers the Raspberry pi to open the attendance database if the validation is successful. Student attendance is taken according to the UIN spoke out by the student and this continues until the last student updates his/her attendance. The server receives UIN and the status and proceeds to update the records present. Alexa hosted on the Echo dot, processes the voice command given to the device. Here, ngrok is used to establish the secure tunnel between Raspberry pi and Echo dot in order to help in communication. This app is run from the Raspberry pi and will provide the URL of the local server. This URL is given to Alexa, which uses this as it's endpoint to transmit the data to the server present at the specified URL. Attendance records are stored in the form of a database using Sqlite3. Python is the programming language used by them to code in developer console. Alexa skill kit helps in creating this skill and is implicitly converted to JSON code by the developer console. Here they use 'pi' as their

invocation name and made three intents for skill called Auth, Ses, Att. Auth, Ses, Att are intents in order to establish teacher’s authentication, get server number and to process UIN of all students present respectively. This system is more effective than the conventional system and efficiency is almost greater than 80%.

2.2. Computer Based Psychotherapy for Treatment of Depression and Anxiety

This paper describes a new computer-based delivery method of Cognitive Behavioral Therapy for relief of depression or anxiety. The system is called Blue Begone. It focuses on alleviating deficiencies like depression and anxiety using computer-based techniques by using the principles of the psychotherapy method called CBT or Cognitive Behavioral Therapy. The therapist evaluates the client, decides on a course of treatment and begins to apply it. As the therapy proceeds, new information may result in a modification of the treatment regimen. This appears to be a simple feedback control loop. In practice there is considerable noise in measurement of the control error. In Blues Begone, the primary control error is the score obtained using a depression measurement instrument. The instrument we use is called the Purves Depression Questionnaire (PDQ). In addition to measuring degree of depression, the initial use of the PDQ includes some questions that deal with particular ancillary factors like alcohol or drug use, bereavement, self-esteem, etc. (Purves) The combination of personal demographic data, PDQ score and additional answers give Blues Begone enough data to prepare an initial Road map to Recovery. At base, Blues Begone includes a collection of 30 episodes. Each episode is not rigidly specified but rather is defined in terms of a scripting language. The scripts accommodate contingencies to respond to PDQ answers, and permit tailoring of the program as it appears to a user. After completion of the first PDQ, the scripts for each episode are executed, and the corresponding initial Road map to Recovery is computed. As the user proceeds through the road map of thirty episodes, the content is likely to be further modified to reflect observed need. User data is maintained in an encrypted database. This is actually an HTML document however it is retrieved strictly from the operations disk using a dedicated browser that appears on the Blues Begone desktop without recourse to the internet. A Talking Head consists of a cartoon head with multiple mouth and eye arrangements synchronized with a speech synthesizer. We use synthetic voices compatible with the SAPI 5 standard. SAPI 5 allows the speech engine to generate events on word, viseme and utterance boundaries, so that the head can, for example, match mouth shape to the sound currently being made. The text to be spoken is in the form of an XML document that passes through a preprocessor before being sent to the speech engine. The pre-processor replaces tags with things like the user’s name, time-of-day-aware greetings, PDQ score, etc. Talking Heads

can bring up other tasks, Information Units or Cartoons. One application uses a Talking Head to narrate a set of charts as they are cycled under script control. This flexibility means that the Talking Head can be quite personal and context aware in talking to the user. The speech script allows one particular talking head to be replaced by another during the speech. Blues Begone is an innovation in delivering self-help to people suffering from depression and anxiety.

3. METHODOLOGY

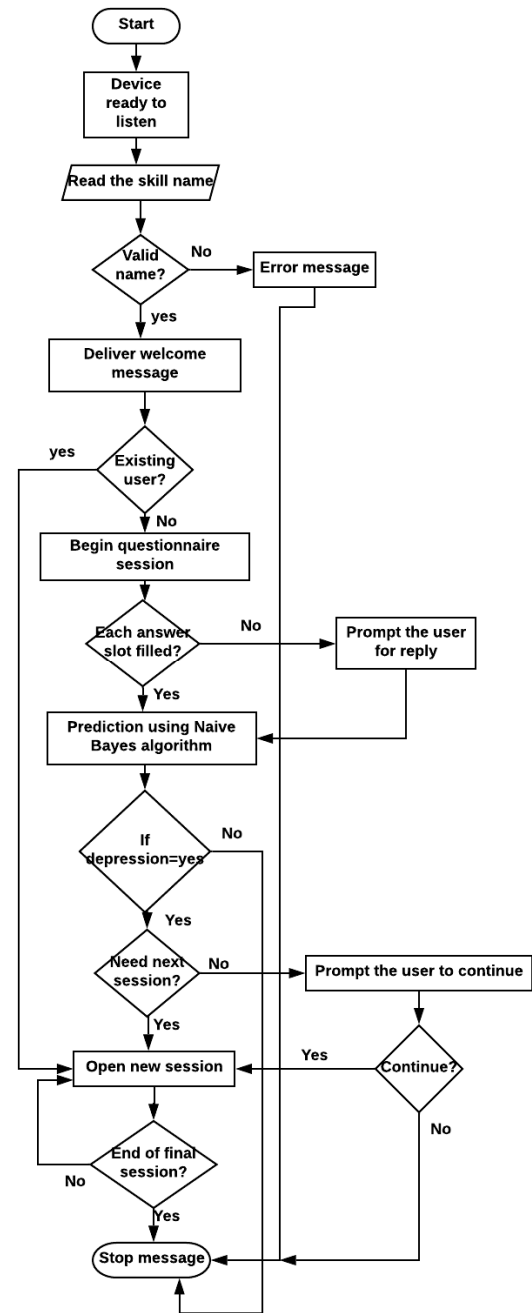


Figure.1:Flowchart

The flow chart for the project can be depicted as:

The working of Therapy AI is a combination of more than one platform or Services. To the users who are using the skill it seems like a single service that they are in conversation with the device. But to view this in a different perspective it is the collaboration of Alexa Developer console, AWS lambda, AWS s3 and AWS Sagemaker. Along with the intermediate stages like SQS, IAM Roles. Basically, when the user interacts with our project the first line of process is done at the VUI of the project where the user is in direct conversation this is done by the Alexa Developer console. The data collected from the user is transferred to AWS lambda, which is the back end of the Skill that we are using. The Lambda function is the main part of the project. It deals with the decision making of responses and also storing of data as well as calling the corresponding functions. Every request invoked by the user will be resolved by this function. Being in a conversation, the lambda function also needs to prompt the user to repeat or fill in the details which are required. Further the lambda function stores the data from the users to a Simple Storage Service also provided by Amazon. This is done using the boto3 module which will be discussed later. These data which are stored can be used any time during the execution of the code. One can also choose the format in which you can store the data. Here the data is stored in JSON format.

4. MODULE-WISE DESCRIPTION

4.1. Alexa Voice Service

Alexa Voice Service is an important module of this project as it provides access to all cloud-based Alexa capabilities. Alexa voice service does this with the help of Alexa Voice Service API's, hardware kits, software tools and documentation. This project focus on a user-device interaction model, and using Alexa voice service, voice recognition and natural language understanding can be made possible. The user speak out to the device and this voice output is converted to text and send to cloud for natural language processing. And then the response from cloud after processing, which is in the form of text is converted back into voice and delivered to the user. In the Alexa Voice Service environment, services are known as skills. 'Therapy AI' is such a skill that we are trying to develop. Alexa Voice Service provides an easy-to-use Software Development Kit (SDK) called AVS Device SDK for commercial device makers to integrate Alexa into connected product. This Software Development Kit provides C++ based libraries which enable the device to process inputs and triggers to establish a persistent connection with Alexa voice service, and manage all Alexa interactions.

4.2. Developer Console

Alexa developer console manages skills in the developer

console. It provides streamlined experience to help you create, manage, and publish skills. The console organizes skill building into some main tasks:

1. Build

Build is used to set up your skill, configure the interaction model and specify the end points for your service. The specific option completed here depend on the model you add to the skill. There are different models available to add to your skill like Custom model, Pre built models for example, Flash briefing model, Music, Video, Smart Home etc. For a custom model you can create intents and sample utterances. Generally, a model comprises of conversation entities such as intents, slots and utterances. Slots are just like variables which can hold specific value. Pre-built models are interaction models that contain a package of intents and utterances that you can add to your skill. The proposed skill 'Therapy AI' is build using custom model thereby we can create intents and sample utterances.

Therapy AI is provided with Intents like MyNameIsIntent to get the name of the user. Also, the questions asked during the first session of the therapy are given as intents. Utterances are the ways users might interact with the skill and ask it to do something. Intents will trigger the appropriate functions in the backend. The use of slots makes it easier to work on. The slots are variables in utterances. These can have pre-defined values but are empty by default. To define a slot, write the variable in curly brackets in the sample utterances. After building your skill, you need to set your end points. This is where Alexa will send user intents to be processed. End points will run the code necessary to interpret and provide responses for user intents and send JSON back to Alexa to turn into speech. Since end point receives and returns JSON you can use any language you want to author to author your end point. And finally build action is performed for the interaction model.

2. Test

Test is used to test your skill with either text or voice. To test the skill Therapy AI, invoke the skill using the invocation name "run therapy session" and the questionnaire session begins.

3. Distribution

This is used to preview how the skill will appear in the skill store.

4. Certification

This is used to validate your skill, run pre-certification test and submit the skill for certification.

5. Analytics

To review metrics for the skill such as number of utterances, customers and intents invoked. There's also an option to

choose a method to host your skill's backend resources. Like python, Node.js, or provision your own in the Developer console.

4.3. AWS Lambda Console

AWS Lambda is a compute service where you can upload code and code will be executed using Amazon's infrastructure. Each piece of code is called function on the lambda platform. Lambda automatically scales the application by a running code in response to each trigger code runs in it and processes each trigger individually, scaling precisely with size of the workload. Scaling here is done automatically based on the size of the workload. Lambda functions can be written in Java, Python or Node.js. All the functions that you upload to lambda here to be stateless. It helps to connect to other AWS services like S3 or DynamoDB to store stateful information. Lambda function can be triggered by events coming from other AWS services. For example, you can trigger a lambda function when a new file is uploaded to an s3 bucket. Lambda function can also be triggered through HTTP calls. This allows to build an API that is very scalable, cost efficient and maintenance free. Skill id is linked with AWS Lambda function and vice versa. Backend code must be able to handle all the intents and give appropriate responses. These will be spoken by Alexa as voice response. Proposed skill "Therapy AI" uses Python as programming language for the backend. Each response has to be carefully tested to handle all the possible inputs.

Python is a robust programming language and provides an easy usage of the code lines, maintenance can be handled easily. We use Python for our project to code on Amazon Developer Console and also at the backend. Each intent in the skill are coded as separate Python functions in the code. We can use other languages to code for the skill, but we selected Python because it is easier to learn and to modify and have an extensive-support libraries.

4.4. Amazon SageMaker

Amazon SageMaker is fully managed service that provides environment for developers to build, train and deploy Machine learning models quickly. It helps to develop high quality models in an easier way by removing heavy lifting from each step of Machine learning process. To get the service of SageMaker for your skill, you can give permission to SageMaker using IAM. (It provides jupyter authoring notebook instance for easy access to the data source for exploration and analysis).

It provides common Machine learning algorithms in optimized form to run efficiently against extremely large data in a distributed environment. By launching model using, Amazon SageMaker, we can deploy it into a secure and scalable environment. This service is paid and is billed by minutes of usage, with minute's fees and no upfront

commitments.

Amazon SageMaker includes many features such as Amazon SageMaker Studio, Amazon SageMaker Ground Truth, Amazon Augmented Artificial Intelligence, Amazon SageMaker Studio Notebooks, Amazon SageMaker Experiments. SageMaker helps to integrate ML to Alexa skills. To train a model, we need an example data, the type of data we need depends on the problem that we want our model to solve. To pre-process the data, initially we have to fetch, clean the data and transform the data to improve the performance. In SageMaker, Jupyter notebooks used on your notebook instance to pre-process the example data. Notebook is used to fetch the dataset, explore it and prepare it for model training. The next step is to train the model. Model training includes both training and evaluating the model. Then the final step is to deploy the model. With Amazon SageMaker, we can deploy the model independently, decoupling it from the application code. Amazon SageMaker processing, along with other critical Machine learning tasks provided by Amazon SageMaker, such as training and hosting, provides the benefits of a fully managed ML environment. After a job is submitted Amazon SageMaker launches the compute instance, processes and analyses the input data and releases the resources up on completion.

Amazon SageMaker provides training algorithms. If any of these meet your needs, it will be a great solution for quick model training. Amazon Debugger can be used to inspect training parameters and data throughout the working. It alerts the user if an error occurs. Then we can submit the custom python code. After the creation of training job, Amazon SageMaker launches the Machine learning compute instances and uses the training code and training dataset to form the model. It saves the training model in the S3 Bucket which is specified by the user. In the proposed skill we use Naive Bayes machine learning algorithm to classify and predict depression of the user. A data set is loaded into Sagemaker via a library function called Panda which is then compared with the data collected from the user and prediction is carried out. The data set is split into Training set and test set.

4.5. S3 Bucket

Amazon simple storage service is an object storage that offers data availability, security and performance. Objects are aggregated in to a bucket. A bucket is like a name space or a database. Customer of all size and industry can use it to store and protect any amount of data for a range of use cases. There are many storage services provided by amazon like Dynamo Db, S3 bucket etc. There is no initialization handshake to establish connections for S3 and they can scale on demand. S3 can store individual objects up to 5TB.As S3 bucket is scalable, we can store large amount of data in s3.

For our project we use S3 bucket. A data set is loaded into S3. Also, the data collected from the user asked during the questionnaire session is saved in s3. Finally, the predicted output is stored in S3.

4.6. Naive Bayes Algorithm

Naive Bayes algorithm is a machine learning algorithm for classification problems. It is based on Bayes' probability theorem. It is primarily used for text classification which involves high dimensional training data sets. A few examples are spam filtration, sentimental analysis, and classifying news articles. It is not only known for its simplicity, but also for its effectiveness. It is fast to build models and make predictions with Naive Bayes algorithm. Naive Bayes algorithm is called "naive because it makes the assumption that the occurrence of a certain feature is independent of occurrence of other features. Naive Thomas Bayes algorithms are principally employed in sentiment analysis, spam filtering, recommendation systems etc. they're quick and simple to implement however their biggest disadvantage is that the need of predictors to be freelance. In most of the important life cases, the predictors are dependent, this hinders the performance of the classifier.

5. RESULTS AND DISCUSSION

5.1 SKILL INTERACTION WITH USER

The Interaction with the user lets the project to learn about the user. The conversations can be used or edited any way we like. We can even prompt the users to say something and also can take the answers and store them. The very first part of the working will be the interaction with the users. The Soul purpose of the first Session is to gather the information from the user to predict whether they have depression or not.

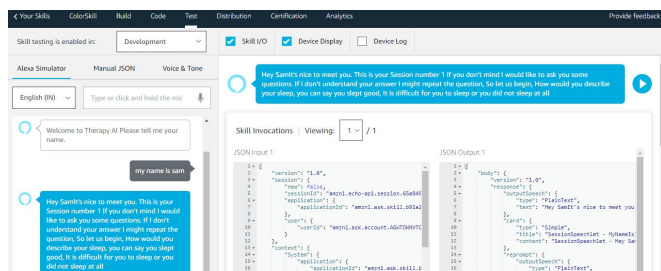


Figure 2: Interaction Output
5.2 S3 Bucket

The data collected from the user and the dataset all are stored in the Simple Storage Service (S3) provided by Amazon. The data to be predicted is also stored here and

can be invoked by AWS Sagemaker, that is the main advantage of this service the fact that all Services can access the same storage space. It is cloud computing at its best. The S3 makes the data ready for further proceedings and calculations.

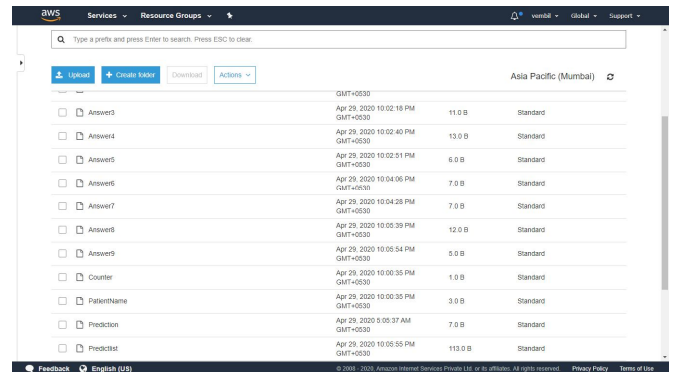


Figure 3:S3 Bucket Data Storage

5.3 Sagemaker

After getting the data from S3 here we have created a model using Naive Bayes Classification Algorithm and it will make a prediction with the received data. The reason why Sagemaker was found best for this project was that it is provided by Amazon and it is easy to integrate with our project. Rather than exporting every data to a separate device for processing we can access it through cloud. The procedure for the prediction is

1. Defining the dataset
2. Encoding the feature
3. Generating the model
4. Prediction based on the model

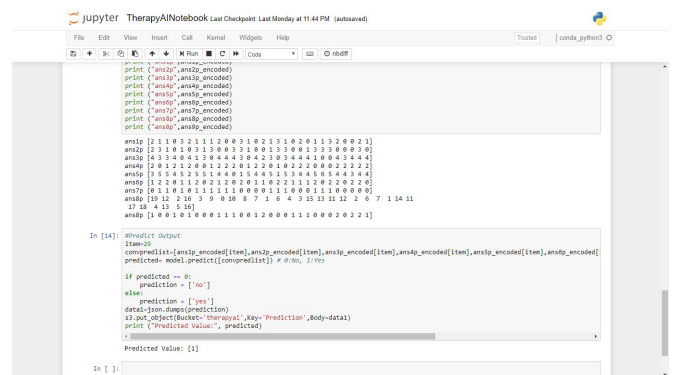


Figure 4: Prediction

5. CONCLUSION

We currently live in a world where Voice Assistants and AI have become a basic or even necessary part of daily life. In developed countries almost every person owns Personal Assistants like Alexa, Siri etc. This is because of the easiness in their usage and hand-free interaction. Even

though the technology has improved to this extent, Mental health of people exists as a big issue. 'Therapy AI' act as a small solution for this issue and it make use of the benefits of Voice Assistant. This Amazon-Alexa based skill will help the people who suffer from depression by providing motivation for them. Amazon provides developer console and various other services to develop a new skill. Machine learning algorithms can be integrated with the skill to make it more effective. Naive Bayes algorithm is used here to predict whether the user is depressed or not. If yes, the user gets suitable motivation. Unlike the real-life Psychologists, this does not charge by the hour and is available to all the people provided they buy an amazon Alexa supporting device. People who are suffering from such conditions must be allowed to be open to someone and this someone might as well be an Artificial Intelligent Assistant where the user and the device engage in a conversation. The project can also encourage people who are afraid or shy to open up to other people. Based on the details given by the user on each session, the user is provided with the best therapy suitable for him. The technology we have today might limit the satisfaction for the user comparing to the satisfaction they can receive from a human but our future has the key to this and this project can be updated with the new features that are invented day by day.

REFERENCES

1. W. P. G. B. Yingxu Wang, ``Perspectives on cognitive computing and applications,`` Inter-national Journal of Software Science and Computational Intelligence, p. 32-44, 2012.
2. Adithya Sanjeev Byalpi, Anush, `` (Anush, 2018)`, Proceedings of the International Conference on Communication and Electronics Systems (ICCES 2018)
3. Byron Purves and David Purves, `` (Purves) Z. Huang, "An investigation of emotion changes from speech," IEEE Proceedings.