Mustansiriyah Journal of Pure and Applied Sciences, Print ISSN: 2957-9910. Online ISSN: 2957-9929, Vol. 2, No. 4, 2024, Pages 157-169
DOI:

MJPAS MUSTANSIRIYAH JOURNAL OF PURE AND APPLIED SCIENCES

Journal homepage: https://mjpas.uomustansiriyah.edu.iq/index.php/mjpas

RESEARCH ARTICLE – COMPUTER SCIENCE

Xeinabot: A Spoken Intelligent Agent Approach Using Hidden Markov Model with Chatbot Technology in Android Environment

Suaad M. Saber [1*], Xeina Abbas [2], Noor H. Muhsin [3]

[1,2,3] Department of Computer Science \ College of Education \ Mustansiriyah University

[*] Corresponding author E-mail: suaad.m.saber@uomustansiriyah.edu.iq

Article Info.

Article history:

Received
11 July 2024

Accepted
1 August 2024

Publishing
30 September 2024

Abstract

The present study describes the designing and its execution of xeinabot, a public domains spoken smart agent in an Android system. The primary goal of this work is to design and construct a personal agent that combines chatbot technology and statistical methods for speech search. The former uses Hidden Markove System Algorithmic procedures and the N-gram technique for Automatically Speech Recognition, while the latter relies on chatbot systems for Speech Recognizing and uses Multinomial Naïve Bays Method for classification and brute force search for matching patterns. The main objective of this strategy is to create an agent that convinces people into thinking they are speaking with a human when, in reality, they are speaking with a machine. utilizing chatbot technology to make advantage. After evaluating roughly 500 phrases on various subjects, the evaluation findings show very good results and, in the majority of circumstances, the agent behaves like a person. The success task (correct voice recognition and speech understanding) amounts to 86.6%.

The official journal published by the College of Education at Mustansiriya University

Keywords: Automatic Speech Recognition; Speaking Intelligent Agent System on Android; N-gram algorithm; Spoken Language Understanding.

1. Introduction

The most common and efficient method of communication is specking, therefore spoken interfaces are crucial for human-machine connection in many situations. For example, speech input is a useful substitute when traditional input methods are challenging or unavailable. "The field of studies centered on the relationships connecting the language of humans and computers" is called natural language processing, or NLP. It is a branch of computer science and artificial intelligence (AI) [[1], [2]]. Additionally, NLP is defined as "software programs that assess, comprehend, or generate multiple forms of spoken language, such as Arabic, Russian, Japanese, or English." Text, speech, or keyboard input are possible input forms [[3], [4]].

The term "an event that operates and reacts within a given environment" is an umbrella description of an agent. People often use portable speaking agents as a common form of communication in their daily lives. That the user can talk to the device and the device listen and answer the request, and then the cellphone is going to be as intelligent as a personal helper [[5]–[7]].

The agent in this paper is work in Android environment This is a phone-specific software framework. Along with customer apps, frameworks, code libraries, and much more, it is with an operating system built on the Linux kernel. in addition to standard phone operations. The operating system's components are developed in the programming languages C or C++, and users can create Android applications using the Java language [[7], [8]].

The intricate design of intelligent agent systems is due to their ability to evaluate and generate speech units, as exemplified in Fig. 1, which depicts the user-system communication cycle during the flight booking process [[9], [10]].

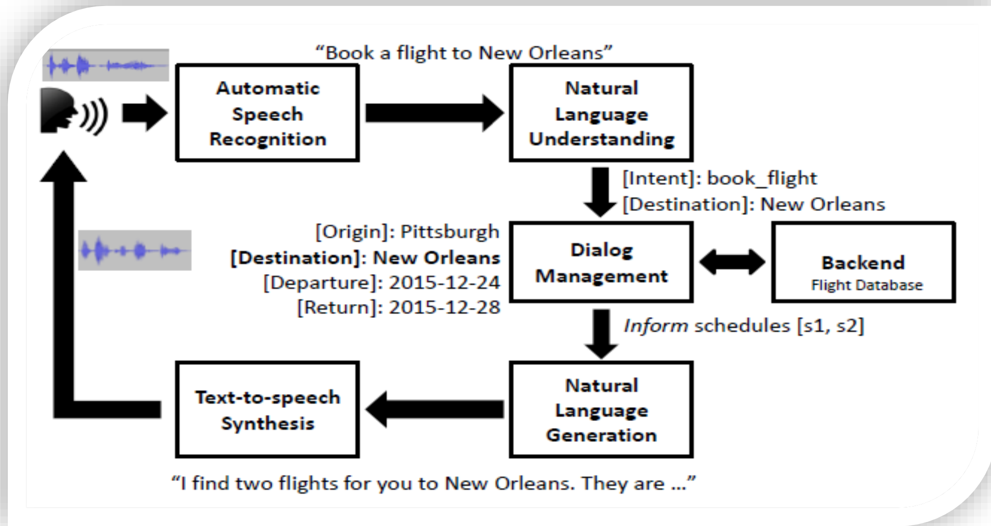


Fig. 1. Spoken intelligent agent system [[11]]

Automatic Speech Recognition (ASR) function is to accept an audio waveform as input and output a string of words as the output. This consists of three stages: Language model, Decoder, and Acoustics Model [[12]–[14]]. The Spoken Language Understanding (SLU) can be accomplished using two primary techniques. Firstly, some of the methods such as Tokenization, tagging, and grammatical rule are used. The dialog manager must then choose what course of action to follow [[15]–[18]]. Secondly, using of chatbot technology, where the goal of programming via Chat-Bot is to have a discussion by applying intelligent or logical ways [19]. The Text-to-Speech Synthesis methods are used for generating spoken speech output from a primary source of information [[20]–[24]].

2. Related works

2.1 ALICE 2007: A.L.I.C.E. (the Artificial Linguistic Internet Computer Entity) is produced by Richard Wallace, training using dialogue corpora was the main aim behind creating ALICE it contain simple patterns templates and simple matching technique .

It has the capacity define partial of the programmers that it processes and to characterize the (AIML objects) type of data object. These objects have two unites (topics and categories); these categories contained data either parsed or unparsed and the bot in pattern matching inputs generated responses:

<pattern> (input) <template> (output)

These pairs stored in a knowledge base as document. An extension of XML, called AI Markup Language (AIML),are used in these documents. ALICE is a (3 time) winner a competition held each year awards the most intelligent chatbot called the Loebner prize, ALICE come with data that exist in its database and does not have the skill to learn .

2.2 Scientific Learning Reading Assistant™ 2010: Scientific Learning's Reading Assistant takes a proven technique for improving reading with feedback and makes it practical to give a classroom or a school full of students this essential practice, as often as they need it. At the same time, Reading Assistant Plus monitors and summarizes students' progress for teachers and administrators.

The most important component of Reading Assistant is speech recognition and speech processing is the speech recognition software, which allows Reading Assistant to listen and follow along as the student reads. Reading Assistant uses the PocketSphinx speech recognizer, which is a part of the CMU Sphinx Toolkit for Speech Recognition. The PocketSphinx version is kept as up-to-date as possible to take advantage of the latest speech recognition features and enhancements

2.3 ILA Voice Assistant 2015: ILA is a voice activated personal assistant very similar to Apple's Siri, Microsoft's Cortana or Google Now, It is designed to run on desktop or media centre PC and integrates into your home environment and it's very interesting for smartphones. ILA can search the web for you, find locations, get directions, start timers and set reminders, read news headlines, open programs, execute system commands and much more, It's especially fun when used with a bluetooth headset or microphone array to freely move around in your home while talking to ILA.

2.4 Google Assistant 2016: Google introduced the Google Assistant, which lets users have a natural conversation with Google. This Assistant is helpful, simple to use, available where user need it . The Hardware that works with this Assistant is Android phones, iPhones, headphones, voice-activated speakers like Google Home and others from several manufacturers, Android Wear and Android TV.

The Smart home devices and platforms can work with this Assistant, it can now control over 1,000 smart home products from more than 100 brands. The assistant introduced Hands-Free Calling, reminders, shopping, shortcuts, step-by-step instructions to millions of recipes, and more. And of course Voice Match, which enables different household members to get personalized help on a shared device.

3. Proposed method

This paper is concerned with the design of the planned intelligent agent and a layout explanation of the designing and implementation. It contains an explanation for each implemented stage related to the applied system.

There are three primary phases to the complicated system described as the Suggested Intelligent Agent System, as shown in Fig. 2.

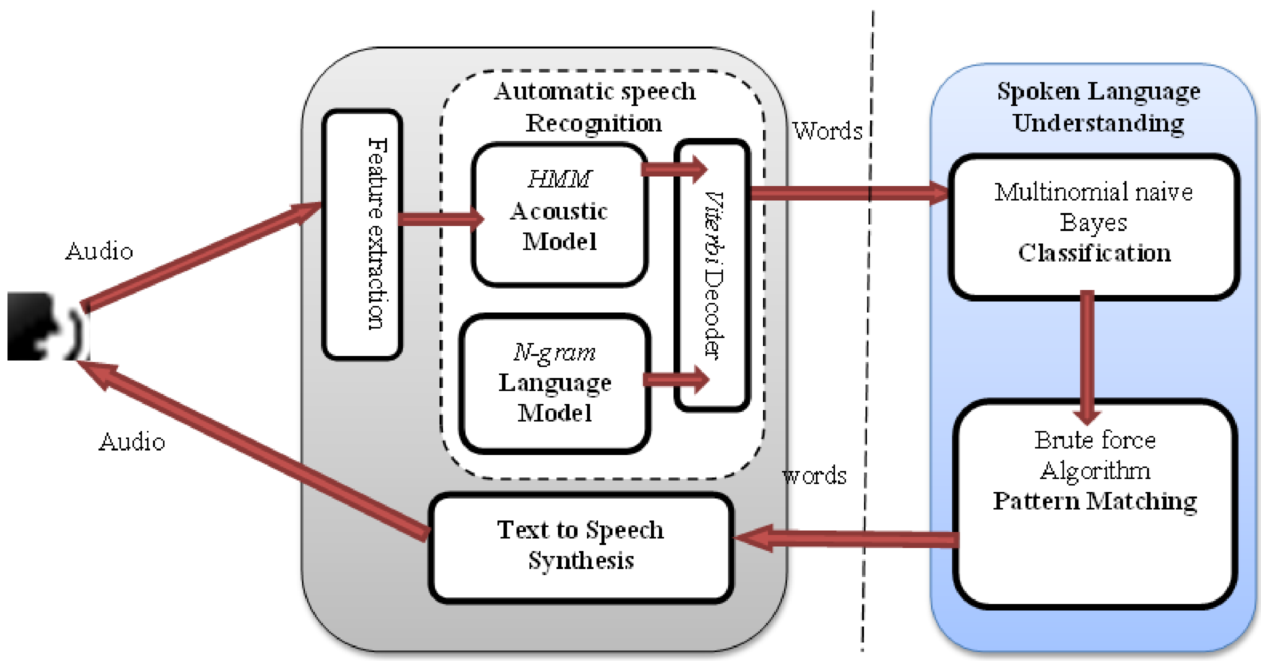


Fig.2. The Proposal System Architecture

- First phase: Automatic Speech Recognition (ASR) using statistical approaches (Forward, Baum welch, Viterbi, N-gram algorithms)
- Second phase: Spoken Language Understanding (SLU) using chatbot technology (Multinomial naive Bayes, Breadth first algorithms)

- Third Phase: Text To Speech (TTS) using Android API.

4. Automatic Speech Recognition

Automatic Speech Recognition (ASR) using statistical approaches (Forward, Baum welch, Viterbi, N-gram algorithms). The ASR contain several algorithms and techniques as shown in the block diagram Fig.3.

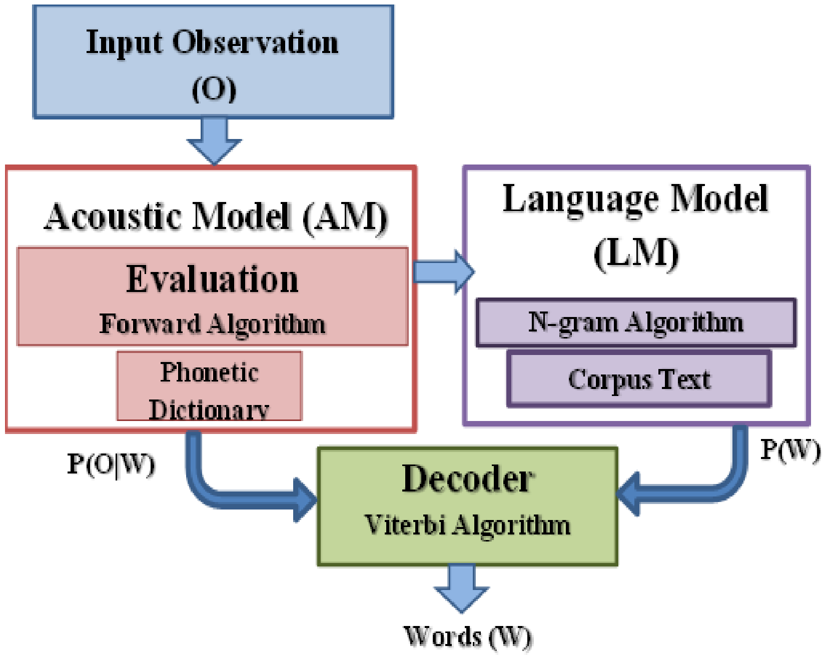


Fig.3. The Automatic Speech Recognition Steps

The following is how to identify speech:

- split the waveform at utterances by silences.

Make an effort to identify the sentence. by matching every word combination that might possibly exist with the audio. Next, select the word combination that matches most closely. This wixll explain the isolated word recognition in a language of three words only {'one', 'two', 'three'} as an example to simplify the process, so for each word to be recognized the steps will be:

a. If will train the possible words in the language to have HMM models for each, that mean the words {'one', 'two', 'three'} will be trained to have three HMM models {M1, M2, M3}, each with three states and three productions. As shown in Fig.4.

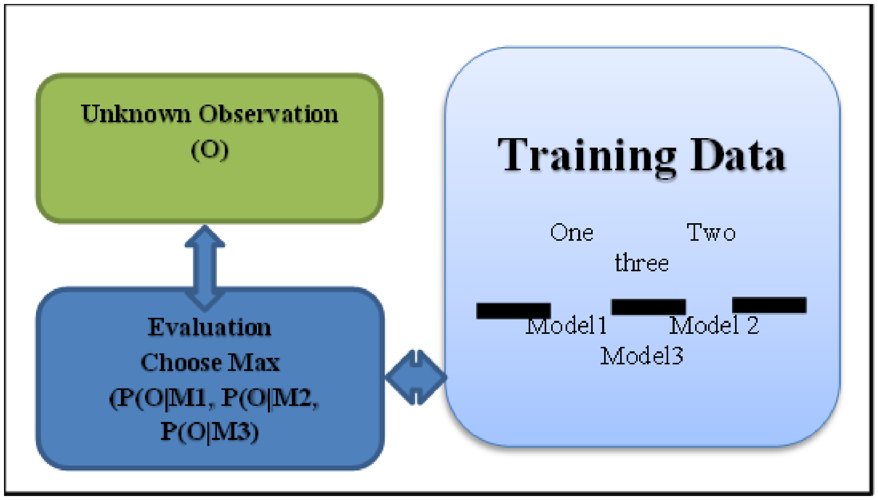


Fig.4. Isolated Word Recognition

b. The training of sample of the specific word that been recorded.
c. Now, to recognize the word, pick a model Mi, that it is the maximum of P(O|Mi), i element of L, (L number of words).

It is also important to mention that the work is related to HMM basic problems, two of these problems are used in the Acoustic model and the third is used in decoding step, the three problems can be summarized in the following:

- Problem 1 Evaluation:
  - The problem: Compute the probability P(O/W).
  - Given: a HMM model M and observation Sequance O .
  - Solution: Using the Forward algorithm.
- Problem 2 Decoding:
  - The problem: finding state sequence which maximizes probability.
  - Given: a HMM model M, and Particular sequence O.
  - Solution: using Viterbi Algorithm.
- Problem 3 Training:
- The problem: Change model M parameters to raise the probability of an observation sequence.
- Given: Observation sequence and Model M.
- Solution: Using Baum-Welch algorithm.

4.1 Feature Extraction

The study uses CMU sphinx toolkit for feature extraction, this tool work on acoustic signal and produce phones as output using silence to separate the words, Mel-Frequency Cepstral Coefficients (MFCCs), which quantify characteristics like frequency and amplitude, provide the basis for this technique. After the audio representation is produced, the input is split up into frames, which typically last 10 milliseconds. A vector of MFCC features is represented by each 10-ms frame.

4.2 The Acoustic Model

The Acoustic model use two type of HMM called Forward, and Baum welch algorithm (which is in fact a Forward Backward Algorithm), this process work on the phones generated by Feature Extraction step comparing it with the phonetic dictionary to find the probability for each word.

a. Phonetic Dictionary: The CMUSphinx (CMUDict07) phonetic dictionary for the dialect of English is used by the system that is suggested to map vocabulary words to phoneme sequences. A new dictionary can be created from scratch or expanded with the use of several tools; one such tool is the g2p-seq2seq tool. It offers cutting-edge conversion accuracy and is built on neural networks and integrated into the Tensorflow framework.

b. Evaluation: The analysis of the probability of an outcome specific sequence is the first stage. which the Forward algorithm may determine well. The probability that any given set of events has resulted in the given set of observations is calculated by the Forward algorithm. A matrix contains the calculations for the various events sequence paths.

Given an HMM model λ = (A,B,π) and an observation sequence O={o1,o2,…,oT}, the Forward algorithm process can be explained as follows: find the probability P(O| λ). By taking into account a single distinct state sequence Q = {s1, s2, s3,..., sT }, where si ∈ S, then calculating it using the observation probability supplied by the state sequence.

In HMMs, any sequence of state can produce the observation. This means that the observation probability is the sum of observation probabilities for all the possible state combinations:

$$P(O|\lambda) = \sum_{all\ Q} P(O|Q,\lambda)P(Q|\lambda)$$

$$= \sum_{all\ s1,s2,..sT} \pi_{s1} b_{s1}(o_1) a_{s1s2} b_{s2}(o_2) a_{s2s3} b_{s3}(o_3) \dots a_{sT-1st} b_{\square}(o_T)$$

The steps can be explained as follows:

Steps	Operations
1. Initialization	$\alpha_1(i) = \pi_i b_i(o_1), 1 \leq i \leq N$
2. Induction	$\alpha_{t+1}(j) = b_j(o_{t+1}) \times \sum_{i=1}^{N} \alpha_t(i) a_{ij}, 1 \leq j \leq N$, and $1 \leq i \leq T-1$
3. Termination	$P(O|\lambda) = \sum_{i=1}^{N} \alpha_T(i)$

4.3 Training

Finding the most likely parameters to model a system given a series of sequences that originated from it is the second challenge with the acoustic model. The majority of solutions simply took into account learning from one sequence at a time rather than the challenge of learning from a series of sequences.

4.4 Language Model

The LM uses the n-gram algorithm to generate sentence, probabilities for words given the observation of the previous (n-3) words.

Restrict word search is used in LM. It defines which word could follow previously recognized words and helps to significantly restrict the matching process by remove words that not acceptable.

This process depend on a number of text files used for training, and because the proposal system is open domain then many text files needed in order to cover the words in the language.

- **a.** Corpus Text: geting an extensive collection of pure texts ready. Extend acronyms, turn numbers into words, and remove non-word objects. In addition to the text files that are freely available online for language modeling, in addition with using Wikipedia XML files that are able to acquired and cleaned using a specialized Java toolkit.
- **b.** N-gram algorithm: Using the corpse text to be an input source, this algorithm determines the optimal word order. Let the set of potential documents be the sample space. Then, introduce two random variables, N and W, where W = (W1,...,WN) is the document's word sequence and N is the document's length, or total number of words. That provides just a probability distribution P(W) as a language model. For a given vector sequence, the search method finds the most likely word sequence among all possibilities W:

$$\text{MAX } \{p(W/O)\} \quad \ldots.(1)$$

Estimating the probability of the following is the aim as well:

$$p(O|W) = \sum_{s,p} p(O,S,P|W) \quad \ldots.(2)$$

A simple example to the process, is using Trigram (n = 3) language model, so the probability of the sentence: “I love my country” is the probability of P(I, love, my, country) which calculated using N-gram as:

P(I|<s>) P(love|<s>, I) P(my|I, love) P(country| love,my){\displaystyle {\begin{aligned}&P({\text{I, saw, the, red, house}})\\\approx {}&P({\text{I}}\mid \langle s\rangle)P({\text{saw}}\mid {\text{I}})P({\text{the}}\mid {\text{saw}})P({\text{red}}\mid {\text{the}})P({\text{house}}\mid {\text{red}})P(\langle /s\rangle \mid {\text{house}})\end{aligned}}}

start-of-sentence typically denoted <s>.

5. Decoding: The Viterbi algorithm

Finding the "highest possible series of words that produced an outcome series" is the second issue. The Viterbi technique makes computing this efficient. To find the algorithm's maximum probability route, a trackback is employed. During the process, the probability of traveling in such a sequence is also calculated.

This process is to find the local maximize and the output is the most likely sentence, after processing the input from AM which depend on acoustic feature of the observation and the input from the LM which depend on the logical sequence of the words in the sentence the Viterbi algorithm find the final result.

By making use of the maximum (or Viterbi) approximation, the optimization problem can be written as:

$$\hat{w}_1^N = argmax_{w_1^N}\{[\prod_{n=1}^{N} p\big(w_n\big|h_{n)}\big]\,.\,max_{s_1^T}\{\prod_{t=1}^{T} p(s_t|s_{t-1},w_1^N).\,p(x_t|s_t,w_1^N)\}\}] \quad ..(3)$$

This simple example to summaries the procedure of the ASR phase, :

The user said: I need two second to finish

The Forward procedure bult a matrix for sentances probabilities:

I need to seconds to finish

Eye need two seconds to finish

I need to seconds two finsh

.

The language Model gives the Forword matrix probabilities:

0.6	0.3	0.7	0.0	0.6	0.3
0.3	0.5	0.6	0.4	0.5	0.7
0.6	0.1	0.5	0.3	0.2	0.3

The final step is the Decoding that choose the higest path from the AM and LM.

6. Spoken Language Understanding (chatbot technology)

Their are three primary stages in this phase. The first is the output from the recognition phase; the input is a sentence, and the second is a text that needs to be read or a command that has to be followed. The primary steps of this phase are depicted in fig.5.

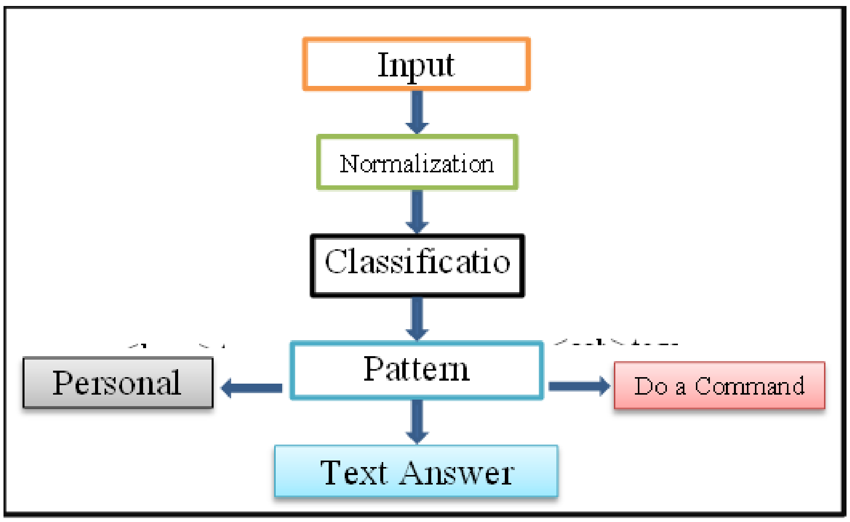


Fig.5. The Speaking Understanding Diagram

6.1 Normalization.

In the work the normalization step do not contain substation or Sentence Splitting or other type of normalization because the sentence is an output from ASR phase. But in checking in this step if there is a relation between the new sentence and the last answer in order to continue in the conversation this called the input path.

6.2 **Classification** (**Multinomial Naive Bayes** algorithm for speech.)

"Multinomial Naive Bayes" is the name of a well-known text classification technique. This is its equation:

$$\hat{P}(t|c) = \frac{T_{ct}+1}{\sum_{t'\in V}(T_{ct'}+1)} = \frac{T_{ct}+1}{(\sum_{t'\in V} T_{ct'})+B'} \quad \text{...(4)}$$

The class with the highest score is the one that is most likely to be included in the input sentence. may count the occurrence of each word in each class given a collection of sentences that each belong to a class and a new input sentence. **The multinomial Naive Bayes algorithm** is a traditional approach used in natural language processing (NLP) and text classification.

Because this classifier treats each word in the sentence it is classifying as having no relationship to any other word, it is "naive" in the sense that it presumes independence between "features." The classifier is aiming to classify a statement, not to figure out what it means.

6.3 **Pattern matching with AIML files (Brute force Algorithm)**.

After finding the class of the sentence the search then will be in a specific AIML file to find the suitable answer, It is considered that each bot has a unique collection of AIML files. The personality of the bot character is defined specifically by this set of AIML files. A number of AIML files available on the internet to increase the bot information.

The last step in the system is a simple Brut force search algorithm used to search in a certain AIML files that related to the category been classified.

7. Experimental Results

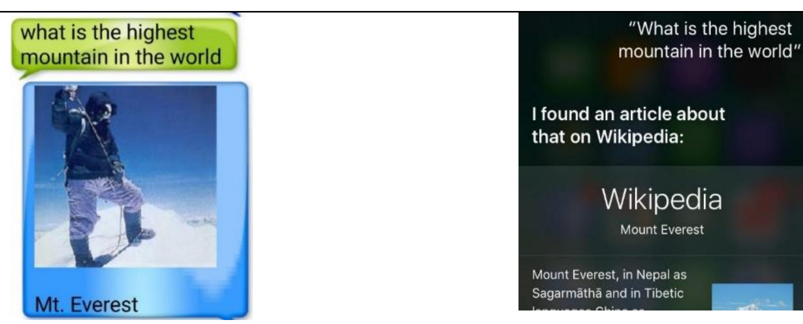
The test consists of 500 questions and an argument of roughly 2,900 words in various subjects that are commonly asked to these kinds of systems. It selected 50 individuals with good English accents to test the agent, and they ask questions like these:

a) Personal questions
b) Scientific questions
c) Emotional questions
d) Command request

In order to test the proposed agent (xeinabot) and Apple Siri agent simultaneously, the test approach is reading the test words to both agents simultaneously and recording the outcome.

The ASR phase, the SU phase, and the entire system are all separately tested. After testing the ASR in continuous speech recognition, then transferring the results to Microsoft Word to identify any potential syntactic or semantic problems.

After posing the query, "What is the highest mountain in the world?" the following success system response is shown in Fig.6. (comparison with Apple Siri):



xeinabot Siri

User: what is the highest mountain in the world?

Fig.6. Sample of success conversation with xeinabot and Apple Siri

After saying the (500) sentence (conversation) to the proposed agent, the results of all answers can be seen in the Table.1. (only the first answer in the conversation counted if it contains more than one turn).

Table.1. Testing Results

Type Of Test	Total Number	xeinabot
The Percentage of success tasks	500	86%
The percentage of word not correctly recognize	2880	1.25%
The recognition rate (words)	2880	98.75%
ASR wrong words	2880	36
Wrong answer	500	53
Syntax error in answer	500	27
Successes Task (acceptable answer)	500	433
Average of turns	500	14%
Number of dialogs	500	70
The minimum number of tern to complete task	500	1
Conversational abilities	-	Yes
Ability to learn from user	-	Yes
Ability to perform android tasks	-	Yes

7.1 Evaluation Criteria

Evaluating the system by using Objective and subjective evaluation.

Fisrt: Objective evaluation

1.Overall system evaluation

The most widely used approach for doing comprehensive system assessment is called PARADIZE. Performance is modeled using this method as a weighted function of the following:

$$\text{Paradise} = (\text{task success}) - \sum (\text{cost of the dialog}) \quad \ldots(5)$$

Task success : the number of success dialog.

Cost of dialog number of: wrong words, wrong answer, wrong spelling, syntax error… Depending on the test table above:

$$\text{Task success in Xeinabot} = 500 - 53 = 447$$

$$\text{User satisfaction (xeinabot)} = (447) - \sum (\text{wrong words} + \text{wrong answer} + \text{syntax error})$$

$$= 447 - (36 + 53 + 27)$$

$$= 447 - 116$$

$$= 331$$

2.Component Evaluation

It is also possible to assess each conversational interface component separately using particular assessment metrics.

a. Automatic Speech Recognition(ASR)

The WER is the main assessment measure for speech-to-text; it is computed by comparing the identified text with a reference (e.g., a transcription by a human expert) utilizing the following equation, where error kinds include: Substitutions (S), deletions (D), and insertions (I), and (N) represents the total number of words:

$$WER = 100\,\frac{(S+D+I)}{N} \quad \ldots(6)$$

Substitutions (S) =29 word, Deletion (D) = 2 words, Insertion (I) = 5 word, Total number of words = 2880

WER= 100 * (29+2+5) /2880

= 1.25%

b. Spoken Language Understanding

Comparing the SLU's result with a reference representations from an experimental set is an assessment process for the SLU component. The following are the most widely used metrics:

- Sentence accuracy: the percentage of correct syntactic or semantic representations.

$$\%f = 100 * \frac{number\ of\ sentences\ correctlly\ represented}{total\ number\ of\ sentances} \quad ..(7)$$

%fc for xeinabot = 100 * (447) / 500 =%89.4

Second: Subjective Evaluation

the program to tested by (50 users) with good English accent and collect their opinions about the proposal agent behavior by choosing (good, satisfied, bad).

Total number of users = 50

The number of users answer (good): 32

The number of user answer (satisfied): 11

The number of user answer (bad): 7

So 86% from the users how use the program where satisfied on its behavior.

5. Conclusions

1) The specking agent with Chatbot technology is efficient and its make the agent look like human, even that need a big number of AIML to be written, but in fact there is a big number of AIML files in defiant subject can be checked and added to the system.
2) The using of HMM deferent algorithms is very efficient depending on the good percentage of write words (98%).
3) It is not complicated approach that can add and teach the agent anything needed.

6. References

[1] A. F. Mahdi, “Face Recognition System Based on Artificial Neural Networks , Commission for Computers and Informatics Institute for Postgraduate Studies,” *M.Sc. thesis*.

[2] A. E. Jacquin, “Image coding based on a fractal theory of iterated contractive image transformations,” *IEEE Trans. Image Process.*, vol. 1, pp. 18–30, 1992, doi: 10.1109/83.128028.

[3] and V. E. B. B. Das, Ripon Patgiri, S. Bandyopadhyay, “Modeling, Simulation and Optimization,” *Springer Nat.*, 2021.

[4] and M. P. B. Martinez, M. F. Valstar, B. Jiang, “Automatic Analysis of Facial Actions: A Survey,” *IEEE Trans. Affect. Comput.*, vol. 1–1, 2017, doi: https://doi.org/10.1109/taffc.

[5] D. M. L. Ballesteros and J. M. A. Moreno, “Highly transparent steganography model of speech signals using Efficient Wavelet Masking,” *Expert Systems with Applications*, vol. 39, no. 10. pp. 9141–9149, 2012. doi: 10.1016/j.eswa.2012.02.066.

[6] T. Hasan and X. Wu, “An Adaptive Fractal Image Compression,” vol. 10, no. 2, pp. 98–108, 2013.

[7] H. Liu, A. Kadir, and X. Sun, “Chaos-based fast colour image encryption scheme with true random number keys from environmental noise,” *IET Image Process.*, vol. 11, no. 5, pp. 324–332, 2017, doi: 10.1049/iet-ipr.2016.0040.

[8] H. Ebrahimpour-komleh, “Fractal Techniques for Face Recognition,” no. August, 2004.

[9] J. Harbi Al-A ’meri and S. M. Saber, “Face Detection Based on Robust Algorithm of Skin Color,” *Int. J. Sci. Eng. Res.*, vol. 7, no. 5, pp. 118–123, 2016, [Online]. Available: http://www.ijser.org

[10] R. Jafri and H. R. Arabnia, “A Survey of Face Recognition Techniques,” *J. Inf. Process. Syst.*, vol. 5, no. 2, pp. 41–68, 2009, doi: 10.3745/jips.2009.5.2.041.

[11] A.-N. Yusra A. Y. and S. D. Chen, “Comparison of Image Quality Assessment: PSNR, HVS, SSIM, UIQI,” *Int. J. Sci. Eng. Res.*, vol. 3, no. 8, pp. 1–5, 2012.

[12] Y. Ben Jemaa, A. Derbel, and A. Ben Jmaa, “2DPCA fractal features and genetic algorithm for efficient face representation and recognition,” *Eurasip J. Inf. Secur.*, vol. 2011, pp. 1–7, 2011, doi: 10.1186/1687-417X-2011-1.

[13] A. Graps, “An Introduction to Wavelets,” *IEEE Comput. Sci. Eng.*, vol. 2, no. 2, pp. 50–61, 1995, doi: 10.1109/99.388960.

[14] Y. Fisher and S. Menlove, “Fractal Encoding with HV Partitions,” *Fractal Image Compression*, pp. 119–136, 1995, doi: 10.1007/978-1-4612-2472-3_6.

[15] D. Der Naturwissenschaften, “Investigations on Search Methods for Speech Recognition using Weighted Finite-State Transducers,” no. April, 2014.

[16] E. Vazquez, R. Mangues, and A. Villaverde, “Functional recruitment for drug delivery through protein-based nanotechnologies,” *Nanomedicine*, vol. 11, no. 11. pp. 1333–1336, 2016. doi: 10.2217/nnm-2016-0090.

[17] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the Inception Architecture for Computer Vision,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 2818–2826, 2016, doi: 10.1109/CVPR.2016.308.

[18] P. Joshi, “Artificial intelligence with Python : build real-world artificial intelligence applications with Python to intelligently interact with the world around you.” p. 423, 2017. [Online]. Available:

https://www.watermarkbooks.com/book/9781786464392

[19] A. Rajadhyaksha and P. Willemen, "Encyclopedia of Indian Cinema," *Encyclopedia of Indian Cinema*. 2014. doi: 10.4324/9781315062549.

[20] G. G. Lee, H. K. Kim, M. Jeong, and J. H. Kim, *Natural Language Dialog Systems and Intelligent Assistants*. 2015. doi: 10.1007/978-3-319-19291-8.

[21] S. Kubrick and A. C. Clarke, "Dave Bowman: Open the pod bay doors, HAL. HAL: I'm sorry Dave, I'm afraid I can't do that.," *SPEECH Lang. Process. An Introd. to Nat. Lang. Process. Comput. Linguist. Speech Recognit.*, pp. 1–18, 2001, [Online]. Available: http://www.cs.colorado.edu/~martin/slp.html

[22] D. Perez-Marin and I. Pascual-Nieto, *Conversational agents and natural language interaction: Techniques and effective practices*. 2011. doi: 10.4018/978-1-60960-617-6.

[23] W. B. Yu and T. T. Ma, "A study on offline handwritten Chinese character recognition based on chaotic iteration," Mustansiriyah journal of pure and Applied Sciences, vol. 2, no. 1, pp. 62–75, 2024.

[24] A. A. Maryoosh and S. Pashazadeh, "Leukemia Detection Using Machine Learning Algorithms : Current Trends and Future Directions (Literature Survey)," Mustansiriyah journal of pure and Applied Sciences,vol. 2, no. 3, pp. 33–49, 2024.