Journal of Theoretical and Applied Information Technology
28th February 2019. Vol.97. No 4

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ANALYSIS OF SPEECH RECOGNITION IN CONVERSATION BOT FOR INTERACTIVE MEDIA USING AUTOMATED LOG

[1]ERI PRASETYO WIBOWO, [2]ALVIAN RAHMAN HANIF AND [3]ARIES MUSLIM

[1,2,3] Gunadarma University

[1]Doctoral Program of Information Technology, Jakarta, Indonesia

[2,3]Faculty of Information Technology and Computer Science, Jakarta, Indonesia

E-mail: [1] eri@staff.gunadarma.ac.id, [2]alvian524@gmail.com, [3]amuslim@staff.gunadarma.ac.id

ABSTRACT

Gunadarma University has been developing a Speech Intelligence Bot used either simply to automate exchanging information or as a better customer services in many industrial needs. For a basic use, this bot should be able to accept people's oral questions and answer them appropriately. This conversation bot is using Sphinx-4 as its base technology and the process consist of human sending speech to speech recognition system and convert it into text, then the text is analyzed with MegaHal Algorithm supported with a knowledge base data, then generates reponse in a text format and convert it into speech that is sent to human. Thus, this research purpose is to develop and analyze the speech conversation bot accuracy with an Automated Logging System, though the analyzation is still attempted manually. Before answering any question, the conversation bot must define sets of arranged words as grammar using JSGF Language to understand any word it might listen. This research has done a test of asking 20 questions with 3 times repetition on each question to accumulate accuracy scores on the recognition system itself and the bot conversation as application. And the result of the test shows that from all generically 90 questions in total was tested on the bot, it shown 90% confidence score in its recognition level. And for the conversation, the bot showed a considerably low confidence score with 60% from all recognized questions are answered correctly. This results might be affected by the lack of environment control, and logic perfection in how the system accepts voices and converts them into processable data, it needs many improvements in speech recognition. Also, a better and faster way to analyze the log data so any further research could achieve more accurate result is suggested.

Keywords: *Speech recognition, Conversation Bot, Logging, MegaHal*

1. INTRODUCTION

Media is the plural form of medium, which (broadly speaking) descrbes any channel of communication. This can include anything from printed paperto digital data. Digital media, which makes up an increasingly vast option of modern communications

Bots are defined as software designed to auotmate things that a person do. The type of bots getting the converage use conversational unit interface and interface with existing messaging application. Bot help the user of the application find information or go the something done in a seamless, automated way through text-based commands.

In North America, the best known example of a messaging application is Facebook Messenger, Whatsapp, also owned by Facebook, features bit integrations, too. There are number of messaging applications and platforms, like Slack, Twitter, etc.

Many researchers has always been eager to deliver better technology for many types of industries. In this case of research, media information is encouraged to be improved with a speech intelligence conversation bot which can do the service like a normal media information center do.

This speech intelligence conversation bot is the combination of speech recognition which can build the ability to listen the user input and convert it to the text and also convert the output text into a voice [1][2], and make it intelligent by chatbot system which can build the ability to understand the user's question and give the right information about it [8].

Journal of Theoretical and Applied Information Technology
28th February 2019. Vol.97. No 4

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

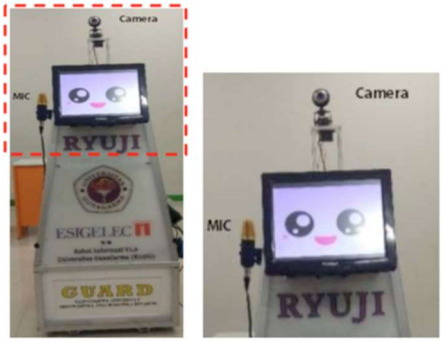

Figure 1: Robot Ryuji

Chat robots (*chatbot*) are developed and researched by many researchers. For example, Siri is the well-known natural speech interface released by Apple[18]. The biggest problem on developing a chat robot is creating a sentence database. Some robot collect from Web data such as Twitter or some SNS. However, these data are including not suitable for enjoying conversation.

After the Chatbot is developed and used by many users, the speech intelligence should be improved so it can understand and give more informations not only for Gunadarma University but also for other type of industries and companies. And this bot application is implied into a robot named Ryuji shown in figure 1[16]. With this mindset, this research will develop a Logging system to saved all conversations from all users and analyze the logs to review, calculate and decide what other options might better to improve the system.

A robot developed by researchers from university of Gunadarma named as RY-UJI as shown in Figure 1, is intergrated with MIC to enable receiving commands and face recognition. This robot is inspired highly inspired by the popularity of biometric data development, which a biometric technology is an individual recognition technology based on special physical features such as fingerprint, hand geometry, iris, and individual faces and behaviors such as voice, hand and handwriting.

This robot is using Eigenface approach, which it is one of the Principal Component Analysis (PCA) algorithm. As this algorithm was seemly the most popular multivariate statistical technique and used by almost all scientific disciplines. Furthermore, this method is able to extract the necessary characteristic data from various type of data by making several inter-correlated quantitative dependent variables as the observation requirements[16].

From the start, this research is meant to be implemented into a robot that is being developed in gunadarma university Thus, following the core development purpose of the robot itself, this research is initially started to fill in a requirement feature needed for the robot, which is to be able to make a conversation with a person. In the end, Ryuji should be able to give a new impression to outside world about gunadarma and be noted as another achievement of gunadarma universty for being able to create a creative innovation.

2. RELATED WORKS

In terms of biometric technology, it has already been studied by many researches using various algorithm and approaches. In general, there are two main steps, which are face detection , face recognition and speech recognition . And the research about speech conversation bot was already proposed by many researcher for many different tasks. Most of the chatbot nodays are largely based on the original chatbot by Joseph Weizenbaum [19]. Mitsuku chatbot created by Steve Worswick is also based on AIMI [19]. AliceTalker [8] was a interactive application which is run by synthezing spoken words and use AliceBot server that is able to run on the same or remote machine. Imran and Shikha [1] also developed a chatterbot which could make the voice function enabled. These researchers are using AIML for chatbot, Microsoft Voice Synthesizer for speech recognition, There are also other researcher used Sphinx as the speech recognition library to develop their own type of application.

The method for developing this speech intelligence conversation is selected by observing those researches of both speech recognition and chatbot development and combine it like the researchers in [1] and [10] did. However, this research will implement MegaHal styling Markov Chain method based on [9] since it was said that that method could make the bot intelligent. And to record the bot conversation, a logging system is applied as a plugin with a readable format that is proposed in this research.

3. PROPOSED METHOD

The Figure 2. describes the general architecture of how the bot understands human speech. The process consist of human sending

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

speech to speech recognition system and convert it into text, then the text is analyzed with MegaHal Algorithm supported with a knowledge base data, then generates reponse in a text format and convert it into speech that is sent to human.

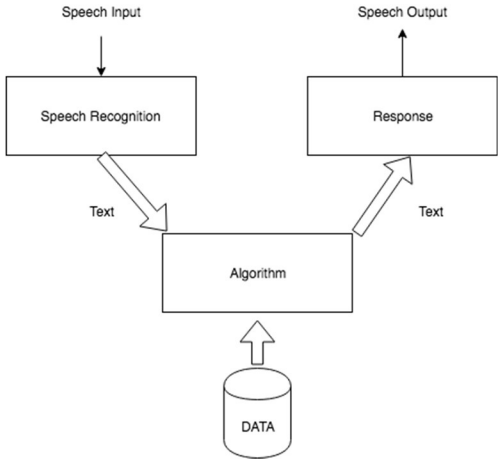

Figure 2: System Architectre

3.1 Sphinx-4

Sphinx-4 is a pluggable framework to help foster new innovations in the core research of hidden Markov model (HMM) recognitions system. It was designed with a high degree of flexibility and modularity. Sphinh-4 was created as the result of inspiration from previous development work on such technology, such as HTK, ISIP, AVCSR and earlier version of the Sphinx system[11].

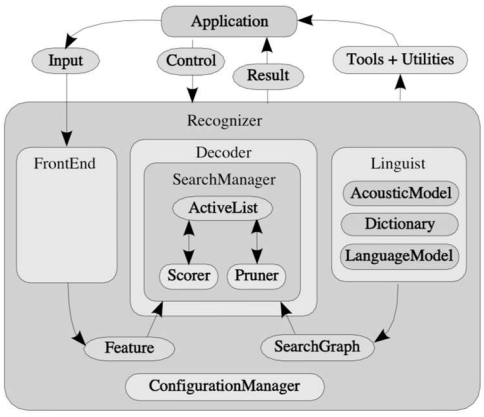

Figure 3: Sphinx-4 Decoder Framework

All of architectures have been shown in figure 3 which consists of various modules that can easily replaced, this allows researchers to experiment with different modules. Sphinx-4 has been provide support for utilities that can support processing level application recognition results, as is the utility included with obtaining grid results, confidence scores, and natural language understanding [10, 11].

Speech recognition in development of speech intelligence conversation bot is using Sphinx-4 library from Java program language.

The Java programming language was originally designed to be used on digital cellular devices, such as mobile phones. When Java version 1.0 was released to the public in 1996. Now its main concentration has shifted to use on the internet, providing interactivity with users by giving developers ways to produce smart web pages. There have been many updates since version 1.0, such as J2SE 1.3 in 2010, J2SE 5.0 in 2004, Java SE 8 in 2014, and Java SE 10 in 2018. Over the years, the Java programming language has developed as a successful language to use on mobile phones and the internet [17].

Java programming built on and improved the ideas of C++ to provide a programming language that was powerfull and simple to use. Java needed to reduce the likelihood of fatal errors from programmer mistakes. With this in mind, object-oriented programming was introduced. When data and its manipulation were packaged together in one place, Java was robust. It was built to include a high level of security. Java is probably the most secure programming language. Programs need to work regardless of the machines they're being executed on. Java language was written to be a portable and cross-platform language that doesn't care about the operating system, hardware, or devices that it's running on.

JavaSpeech Grammar Format (JSGF) defines an independent platform-vendor way to describe one type of grammar, grammar rules. It uses textual representations that can be read and edited by developers and computers, and can be included in the source code. Grammar rules determine the type of speech the user utters (speech disorders similar to written sentences). For example, simple window control grammar might listen to "open a file", "close the window", and similar commands. A grammar is composed of a set of rules that together define what may be spoken. Rules are combinations of speakable text and references to other rules. Each rule has a unique rule name. A reference to a rule is represented by

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

the rule's name in surrounding (<>) characters (less-than and greater-than).

Sphinx-4 consisting three primary modules: the FrontEnd, the Decoder, and the Linguist. FrontEnd takes multiple signals as the input and parameterizes them into a set of features. Then, Linguist made a translation on each type of standard language model existed, along with pronunciation information from the Dictionary and structural information from several sets of AcousticModels, into a SearchGraph. Then, SearchManager within the Decoder uses features generated and SearchGraph from Linguist to perform the actual decoding, and generate Results. The application can issue Controls to each of the modules at any time including during the process of recognition. There are basically three models required for speech recognition in Sphinx4 :

1. Acoustic Model
   This provides a mapping between a unit of speech and can be scored against incoming features provided by the FrontEnd.
2. Phonetic Dictionary (File ends with .dict extension) .
   This provides pronounciations for words found in the Language Model. The pronounciation break words into sequences of sub-word units found in the AcousticModel.
3. Language Model (File ends with .lm extension) This provides word-level language structure, which can be represented by any number of pluggable implementations. These implementation typically fall into one of two categories: graph-driven grammars and stochastic N-Gram models.

3.2 Speech Recognition

The Large Vocabulary Continuous Voice Recognition System (LVCSR) shows a high level of accuracy in recognition in clean conditions from noise or at high signal-to-noise ratios (SNRs) [20].

Speech recognition systems give the computer the ability to listen to users and determine what is said. Current technology does not yet support universal speech recognition: the ability to listen to speech in any context and copy it accurately. To achieve recognition accuracy and reasonable response times, current speech recognition limits what they listen to using grammar.Speech recognition process or framework in general view is shown in figure 4 [17],

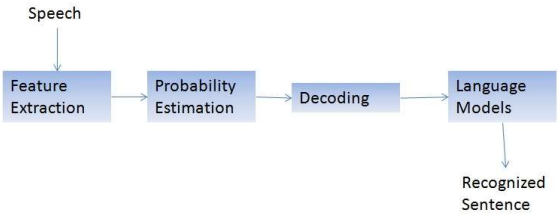

Figure 4. The Speech recognition process

Sequence of words and is determined by integrating many local matches.

Physically, the recorded spoken sentences consisting of acoustic environment and transduction equipment (such as microphone, preamplifier, filtering, and A/D converter), which then generated speech representations will be affected. Any additive noise, microphone position may also be affected.

First, a feature extraction is done to derive acoustic representations that is good for separating classes of speech sounds and is effective to suppressing irrelevant sources of variation. Next, patterns of the core acoustic matching operations is used as the speech recognition. Calculation of the probability estimation is made, between speech frames and spectra frames that used as a data trainer is compared. It can be viewed as a global match to the decoding component. The best sequence of words is searched and determined by integrating many local matches.

Finally, it is the language model, which is use to determine the hypotheses within the global search. This can be further processed as the global decoder output. If the decoding generates more than one sentence, re-scoring sequence on the senteces could be done accordingly to grammar or semantics by language model.

Due to researches[14], speech recognition has a multilevel pattern in term of recognition task, which are examined and structured into a hierarchy of sub words units (phonemes), words, phrases, and sentences.

3.3 Chatting System

The chatbot system that is used in the system is using MegaHal algorithm whish it is one of the method that was published in Loebner Prize for chatbot development[15]. The pupose in making this chatbot is to be able to interpret user words into an understandable information in the system generally by processing any sentences sent by user into a set of data. The step of process shown in figure. 6.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

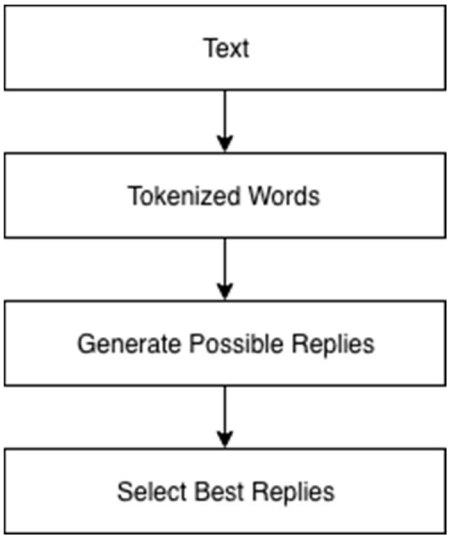

Figure 6: Chatting System

The process is divided into proposed 4 steps:

1. Voices is delivered from human to the Bot as the Speech input. Then processed by MegaHal algorithm into two types: words category and non-words category. Every alphanumeric characters are the words category, while the others are non-words category.
2. The spoken sentence is tokenized into separated words, then will be used as a variable to match any possible replies that will be generated in further process.
3. The system searched any replies based on the provided Knowledge Base.
4. The System chooses the best reply based on counted match between tokenized word used and Knowledge Base data provided.

The MegaHal library used in this research is JMegaHal, the library that designed by Trejkaz [12]. However, for implementing the MegaHal style for this speech intelligence conversation bot, some modifications are applied in JMegaHal library. Those modifications are:

1. Symbol Data Type.

   The data type for symbol implementation is the custom data type. This custom data type contains of start identifier, symbol text, keyword value, and an end identifier. There are two attributes added in this symbol data type, the list of head and the list of tail.

2. Markov Modeling

   Traditionally, the "language" model only captures the sequence of words from a language. The same HMM-based model can be used in two modes: to restore hidden structures such as sentence boundaries, for example, or to evaluate speech recognition hypotheses, thereby integrating sentences into the recognition process [21]. The Markov models build in this research is only from the word category which also categorized as a keyword. The head list attribute function is to list all non-keyword from word category and non-words category that written in before the current keyword symbol until the previous keyword symbol while the tail list function is to list them that written in after the current keyword symbol until the next keyword symbol.

3. Calculate Information

   The modification in this research is deleting the keyword redundancy in calculation of information for each candidate replies.

4. Generate The Sentence Reply

   The chain selected is a sequence of keywords. It needs to add many words for making it readable for human. The process to do this task can be done with finding the similar element in each head and tail list in every keywords in the selected chain.

3.4 Generate The Response

The output of chatbot system is a sentence as a reply for the user that already in readable text form. Since the speech intelligence conversation bot interacts with the user in spoken language so the output must be in speech form. The text-to-speech (TTS) process is done for this task.

TTS is a application for speech synthesis that converts a spoken sound into text. TTS may enable creating a conservative computer, or may simply be used to augment a message. This research is using FreeTTS library to enable speech synthesis and deliver spoken answers as a result. It is based upon Flite; a small run-time speech synthesis engine developed at Carnegie Mellon University with Java programming language.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3.5 Chat Log

In each conversation made, at specifically chosen point of process, which are Input, Process, Output, are recorded into a file. The proposed concept shown in Fig. 7.

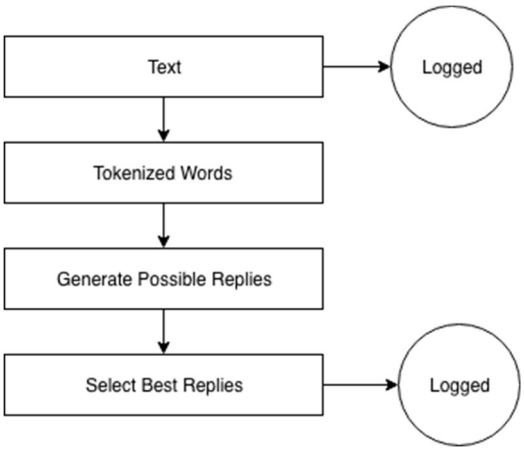

Figure 7: Log in Chatting System

Generally, the system will printed its current processing time and value of the input/output depending on the process step that are currently running. As for example in Fig. 8.

```
[23:07:43]:::recognized:::what are
[23:07:43]:::command-type:::what
[23:07:43]:::response::: gunadarma
```

Figure 8: Conversation Log Example

The file of log is saved by following the current date of time when the conversation is happening. As a result, The chat system will have an organized logs sorted daily.

3.6 Implementation Design

In th is research, the application is technically made into object classes and enumerations. So, to describe the integration form each class with the other, UML Class diagram is made. The Class Diagram is broken down into packages for each component and the relation with other component is made afterwards.

The following are type of packages made in the application.

1. Main. This package is working as the runner of the application.
2. Dictionary. This package is working as grammar collection the conversation bot uses to understand sentences it accept.
3. Response Collection. This package contains several matches between the questions and its related source of answers that the coversation bot use as the output. The source may come as an technical action outside of the application itself, it may be an answer from the data knowledge base it was provided.
4. Library. This package contains all sub programs needed in order to make the conversation bot possible.

To make the conversation bot able to understand every question it should be listen ing. This research added several words including how it should be read. For example, the word “GU NADARMA” is read as in “G UW N AH D AA R M AH”. This is neccesary since it is one of the requirements from the library Sphinx-4 to be able to read any word.

After adding all words the conversation bot needed, this research must setup the sentence sequences that is known as grammar. This will make the conversation bot able to understand the question it listen and able to generate the answer based on the data provided. Here is the example.

```
1 <basic_question>=(
2 what((is|are)[the]visionofgunadarma)|
3 who(areyou|[(is|are)][the]rectorofgunadarma)
3 );
```

As shown, a variable named “ basic_question ” contains 3 possible sentences start from “What” and “Who”, which could be used optionally depending on what speech the conversation accepted at the time. To be exact, the grammar format shown could generate these questions :

1. What is vision of gunadarma
2. What are vision of gunadarma
3. What is the vision of gunadarma
4. What are the vision of gunadarma
5. Who are you
6. Who is rector of gunadarma
7. Who are rector of gunadarma
8. Who are the rector of gunadarma

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4 TEST RESULTS

The speech recognition test that is conducted in this research is based on the vocal range when the user deliver any question as the input. There are three vocal range in this test, according to The New Harvard Dictionary of Music[22].

1. Lower Note. The user is using a heavier voice then usual. In terms of music, it is the notes of Alto F3-D5 and/or Bass E2-C4.
2. Normal Note. The user is using the usual voice, without any attempted change. In terms of music, it is the notes of Mezzo Soprano A3-F5 and/or Baritone G2-E4.
3. High Note. The user is using a higher voice, approaching to a screaming sound. In terms of music, it is the notes of Soprano C4-A5 and/or Tenor B2-G4.

Lower Note and High Note are relatively resulting low accurancies respectively at 45.45% and 18.18% in getting recognized, while Normal Note is showing significantly better accurancy at 77.28%. In this comparison, the result is quite accordance with the hypothesis of what might happen in this test, although it has yet to result a considerably high result (80%-100%). However, not all questions are correctly recognized following with wha the user are actually asking. When using Lower Note and Normal Note, there are 1 to 3 questions that are recognized as another question and resulting respectively 31.81% and 63.64% in total that are correctly recognized. But since the conversation bot is using JSGF Language to help precisely interpret any word in every sentence delivered by the user and is provided with knowledge base data which are tagged appropiately, hence every recognized question are able to get a response which is the answer of what is being asked.

The test purpose of speech intelligence conversation bot is to measure accuracy of the speech recognition and its response recognition by using data provided from proposed Log System. The test of conversation is done by asking 20 questions and repeating each question 3 times. As a result, a file consisting recognized speech and chosen response are made as shown in Figure 9 (*Test Result Conversation Log).*

The questions that was used in the test shown in Table 1.

Table 1: List of questions and the responded answer

No	Questions	Correct Answer
1	What is the vision of gunadarma	Gunadarma will be a leading private university which has international reputations, global networks, and provides significant contributions toward nation competitiveness
2	Who are you	im the information robot from gunadarma university
3	Who is the rector of gunadarma	Professor margianti
4	What is the meaning of the name gunadarm university	Sincere intention of dedication to the society through a high school
5	What defines gunadarma	Defined as a sincere intention of dedication to the society
6	What is the mission of gunadarma	To establish a qualified ict-based higher education to enhance nation competitiveness
7	What are gunadarma student activities	Generally the student affairs activities facilitate interest and student potentialities in their self-development, capacity and talent, art and sport, and social service.
8	When did gunadarm university got ‘listed’	Gunadarma got a listed status by department of education in eighty five
9	When did gunadarma university got ‘equalized’	Gunadarma got an equalized status by department of education in eighty nine
10	Where was gunadarma first located	First located in salemba raya jakarta
11	How did	First located in

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

	gunadarma firstly located	salemba raya Jakarta with two high school
12	How did gunadarma name chosen	As an inspirational thoughts of an architect that built Borobudur temple
13	How did gunadarma built	The gunadarma foundation build two high school then merged into an university
14	How did gunadarma first founded	Founded as gunadarma computer and informatics management high school
15	Find alvian	Looking for alvian
16	Find anna	Looking for anna
17	Find jefri	Looking for jefri
18	Find missmargianti	Looking for missmargianti
19	Find misterdennis	Looking for misterdennis
20	Find mistermusa	Looking for mistermusa

By comparing the data log resulted from the tests and its actual answer described in Table 1, this research counting the accuracy by making percentage after deviding the count of many correctly answered with the time of testing that was done, which is 3 times.

$$\text{Result} = \frac{succeded}{total} x100\%$$

Table 2: Result of tests

No	Question	Recognized	Answered Correctly
1	What is The vision of gunadarma university	100%	100%
2	Who are you	100%	100%
3	Who is the rector of gunadarma	100%	0%
4	What is the meaning of the name gunadarma university	100%	100%
5	What defines gunadarma	100%	100%
6	What is the mission of gunadarma	100%	0%
7	What are gunadarma student activities	100%	100%
8	When did gunadarma university got 'listed'	70%	70%
9	When did gunadarma university got 'equalized'	100%	0%
10	Where was gunadarma first located	100%	100%
11	How did gunadarma firstly located	100%	100%
12	How did gunadarma name chosen	70%	30%
13	How did gunadarma built	100%	100%
14	How did gunadarma first founded	100%	100%
15	Find alvian	100%	100%
16	Find anna	100%	0%
17	Find jefri	100%	100%
18	Find missmargianti	100%	100%
19	Find misterdennis	70%	0%
20	Find mistermusa	100%	100%

Based in the result of tests, it is shown that 90% of all questions are recognized throughout the tests and was answered correctly at a total of 60% accuracy score. Although it seemingly a good result, the bot is far from an actual applicable use, since the accuracy confidence is still considered to be low.

5 CONCLUSIONS

The speech intelligence conversation bot is an application for the implementation of University's customer service improvement. The main purpose of the bot is to provide a better way

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

to deliver information to customer, mainly by listening user questions, analyze the question into the system, and answer it correctly based provided Knowledge Based data.

This might be affected by the lack of environment control, and logic perfection in how the system accepts voices and converts them into processable data, it needs many improvements in speech recognition. Also, a better and faster way to analyze the log data so any further research could achieve more accurate result is suggested.

REFRENCES:

[1] Paul, D. B.," Speech recognition using hidden markov models". *The Lincoln Laboratory Journal*, 1990.

[2] Bansal S., & Bahety, R.," Speech recognition system. Transform the PCM digital audio into a better acoustic representation and Unit matching system", 2013. Available: https://cseweb.ucsd.edu/classes/fa06/cse237a/finalproj/ruchi.pdf

[3] Paul Lamere, Philip Kwok, William Walker, Evandro B. Gouvea, Rita Singh, Bhiksha Raj and Peter Wolf, "Design of the CMU Sphinx-4 Decoder", *8th European Conference on Speech Communication and Technology, EUROSPEECH 2003* , *INTERSPEECH 2003*, Geneva, Switzerland, September 1-4, 2003.

[4] Rashid Khan, Anik Das, " Build Better Chatbots, A Complete Guide to Getting Started with Chatbots", *Book*, 2017.

[5] Amir Shevat, " Designing Bots: Creating Conversational Experiences" , *Book*, 2017.

[6] Joseph Picone, " Java Speech Grammar Format", 2000, *Speech Works International,* Available: http://java.coe.psu.ac.th/Extension/JavaSpeech1.0/JSGF.pdf

[7] Breandan Considine, " Speech Recognition In Java", 2011. Available: https://docs.huihoo.com/javaone/2015/CON7705-Speech-Recognition-in-Java.pdf

[8] Santangelo, A., Augello, A., Gentile, A., Pilato, G., & Gaglio, S., "A Chat-bot based Multimodal Virtual Guide for Cultural Heritage Tours". 2006, *In PSC*.

[9] Liddy, E. D., "Encyclopedia of Library and Information Science", *2nd Edition, chapter Natural language Processing. Marcel Decker, Inc, 2001.*

[10] Trejkaz, "Java Implementation of MegaHal", Available: https://github.com/trejkaz/megahal

[11] Willie Walker et.al, "Sphinx-4: A Flexible Open Source Framework for Speech Recognition", *Technical Report*, 2004.

[12] Anuja Jadhav, Arvind P., " Android Speech to Text Converter for SMS Apllication", *IOSR Journal of Engineering*, Mar. 2012, Vol. 2(3) pp: 420-423.

[13] Manoj Kumar Sharma and Omendri kumari. "Speech Recognition: A Review", *Special Conference Issue: National Conference on Cloud Computing & Big Data,* 2015.

[14] Ajay K.Yadav, Alokit Upadhyay, Shekhar Rastogi and Sonali Shukla, " Controling Of Devices Through Voice Recognition Using MATLAB", *International Journal Of Electrical And Electronics Engineers* , 2015, Volume 07, Issue 01, Jan-June 2015.

[15] Jefri Yushendri, Alvian Rahman Hanif, Anneke Annassia Putri Siswadi, Purnawarman Musa, Tubagus Maulana Kusuma and Eri Prasetyo Wibowo, " A Speech Intelligence Conversation Bot for Interactive Media Information" , *2017 Second International Conference on Informatics and Computing (ICIC)*, 1-3 Nov. 2017.

[16] Anissa Lintang Ramadhani, Purnawarman Musa and Eri Prasetyo Wibowo," Human Face Recognition Application Using PCA and Eigenface Approach, 2017", *2017 Second International Conference on Informatics and Computing (ICIC)*, 1-3 Nov. 2017.

[17] Morched Serbali, Mu'tasem Jarrah and Mohd Taib Wahid, "A Review of Speech Recognition with Sphinx Engine in Language Dettection", *Journal of Theoretical and Applied Information Technology,* 30th June 2012. Vol. 40 No.2.

[18] Eri Sato-Shimokawara , Yoko Shinoda, Tomoya Takatani , Haeyeon Lee , Kazuyoshi Wada and Toru Yamaguchi, "Analysis of Category Estimation for cloud based chat robot", *25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, August 26-31, 2016, Colombia University, NY, USA.

[19] Ridong jiang, rafael E. Banch, "Towards Improving the Performance of Chat Oriented Dialogue System", *2017 International Conference on Asian Language Processing (IALP)*, 2017.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[20] Vikramjit Mitra, Horacio Franco, Martin Graciarena And Dimitra Vergyri," Medium-Duration Modulation Cepstral Feature For robust Speech recognition", *2014 IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)*, 2014.

[21] Andreas Stolcke and Elizabeth Shriberg," Markovian Combination of Language and Prosodic Models for Better Speech Understanding ", *IEEE Workshop on Automatic Speech Recognition and Understanding, ASRU '01*, 2001. Madonna di Campiglio, Italy, Italy.

[22] Angela Oswalt, "Vocal ranges according to the new harvard dictionary of music", 2003, https://web.library.yale.edu/cataloging/music/vocal-ranges.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

```
[22:59:00]:::recognized:::
[22:59:00]:::command-type:::
[22:59:00]:::response:::I can't hear what you said
[22:59:00]:action:
[22:59:02]:::recognized:::
[22:59:02]:::command-type:::
[22:59:02]:::response:::I can't hear what you said
[22:59:02]:action:
[22:59:07]:::recognized:::find missmargianti
[22:59:07]:::command-type:::find missmargianti
[22:59:10]:::response:::looking for information about missmargianti

[22:59:10]:action:findpeople:missmargianti
[22:59:13]:::recognized:::find misterdennis
[22:59:13]:::command-type:::find misterdennis
[22:59:17]:::response:::looking for information about misterdennis

[22:59:17]:action:findpeople:misterdennis
[22:59:19]:::recognized:::find misterdennis
[22:59:19]:::command-type:::find misterdennis
[22:59:22]:::response:::looking for information about misterdennis

[22:59:22]:action:findpeople:misterdennis
[22:59:24]:::recognized:::find misterdennis
[22:59:24]:::command-type:::find misterdennis
[22:59:28]:::response:::looking for information about misterdennis

[22:59:28]:action:findpeople:misterdennis
[22:59:30]:::recognized:::find misterery
[22:59:30]:::command-type:::find misterery
[22:59:33]:::response:::looking for information about misterery

[22:59:33]:action:findpeople:misterery
[22:59:36]:::recognized:::find misterery
[22:59:36]:::command-type:::find misterery
[22:59:39]:::response:::looking for information about misterery

[22:59:39]:action:findpeople:misterery
[22:59:42]:::recognized:::find mistermaulana
[22:59:42]:::command-type:::find mistermaulana
[22:59:45]:::response:::looking for information about mistermaulana

[22:59:45]:action:findpeople:mistermaulana
[22:59:48]:::recognized:::find jefri
[22:59:48]:::command-type:::find jefri
[22:59:51]:::response:::looking for information about jefri
[22:59:51]:action:findpeople:jefri
[22:59:54]:::recognized:::find mistermaulana
[22:59:54]:::command-type:::find mistermaulana
[22:59:57]:::response:::looking for information about mistermaulana
[22:59:57]:action:findpeople:mistermaulana
[23:00:00]:::recognized:::find mistermusa
[23:00:00]:::command-type:::find mistermusa
[23:00:03]:::response:::looking for information about mistermusa
[23:00:03]:action:findpeople:mistermusa
[23:00:05]:::recognized:::find mistermusa
[23:00:05]:::command-type:::find mistermusa
[23:00:08]:::response:::looking for information about mistermusa
[23:00:08]:action:findpeople:mistermusa
[23:00:13]:::recognized:::
[23:00:13]:::command-type:::
[23:00:13]:::response:::I can't hear what you said
[23:00:13]:action:
[23:00:19]:::recognized:::what is vision of gunadarma university
[23:00:19]:::command-type:::what
[23:00:19]:::response::: gunadarma offer Acadmeic program such as Professional Study, Bachelor and Postgraduate
program
```

Figure 9: Test Result Conversation Log