ALICE Pattern Matching Based Chatbot for Natural Language Communication: System Development and Testing

Munaveer Karimeni Peedikayil Tharammal*[1], Muhammad Nauman Bashir[1], Kamaludin Mohamed Bin Yusof[2], Sameera Iqbal[3]

[1]Department of Electronics and Communications Engineering, Middle East College, Muscat, Oman
[2]School of Electrical Engineering, Faculty of Engineering, University Technology Malaysia, Johor
[3]Higher College of Technology

_____

Abstract - The simulation of human processes using Artificial Intelligence (AI) features is finding place in various fields including the e-learning environment. Chatbot system using text and voice, recognizing images, analyzing the sentiments, and generating natural language is latest utility based on AI concepts. The chatbot systems are becoming more common due to their benefits as a support mechanism helping human beings in their day to day tasks. This project aimed at developing and testing a chatbot system to communicate with the education sector users using natural language through android based application creating a smart education environment. The project developed an online chatting platform using Artificial Linguistic Internet Computer Entity (ALICE) Pattern Matching techniques where students could communicate related to their learning activities such as submission deadline of the reports and assignments, student's information, co-curricular and extracurricular activities. The system design uses a Raspberry Pi 3 which works as Transmission Control Protocol (TCP) server and uses three different types of pattern matching techniques which are keyword detection, symbolic reduction, and synonyms resolution. An android user interface application is also designed which works as TCP client. The system design uses a database for student information system in the Python server. The obtained results are in voice and text format from the android user interface application and are displayed on Python Interpreter. The developed project system can enhance the student engagement in learning activities. This system can also help teachers in saving their time and to support them to utilize their class timings for other co-curricular activities like synchronous and asynchronous activities to support active learning and flipped learning. The proposed system has the potential to test and analyze various factors as use of technology, student learning including impact of student engagements in their learning activities.

Keywords: e-Learning; ALICE Pattern Matching; chatbot; AI in Education; Smart Education

_____

INTRODUCTION

In recent years, Artificial Intelligent (AI) systems are one of the advanced technologies which are being used in the field of education, healthcare, industries, etc. Artificial Intelligent systems can simulate the human brain processes using machines (Tim, 2019). These systems are intelligent as humans, so they can interact or communicate with humans through texts or voices. In Tim's article (2019) entitled 'Can a computer fool you into thinking it is human?' published in BBC News states that the chatbots are now present everywhere and handling a large number of complaints and inquiries. Recognitions of images and speech, chatbots, sentiment analysis, and generation of natural language are the applications of AI. AI-based chatbot systems are used for performing day to day tasks, like home automation and

_____

*Corresponding author:
  Email: munaveer00@gmail.com (M.K.P. Tharammal), mbashir@mec.edu.om M.N. Bashir), kamalmy@utm.my (K.M.B. Yusof),
     sameeraiqbal786@hotmail.com (S. Iqbal)

transportation, etc. The chatbot system is a software program or application which communicates with the users using natural language through an android mobile application (Tim, 2019).

In education sector, if the chatbot system is introduced for the students they can know about the due dates of all online report submission for the subjects and their overall education progress and performances. Also, it may help to find information on important events such as exams and extracurricular activities just by interacting with the developed chatbot system. Thus, the developed chatbot system can enhance engagement in all subjects and learning process. The developed chatbot system is powered by Python Transmission Control Protocol (TCP) server and uses three different types of pattern matching techniques which are keyword detection, symbolic reduction, and synonyms resolution. The developer designed and developed an android based mobile application for enabling the use of a chatbot system and a database for student information system in the Python server. The rest of the paper is categorized as follows:

The 2nd section of this paper explores the related work. The 3rd section discusses the methodology applied in order to execute the proposed system. The 4th section reviews the results obtained while the 5th section sums up the results of this project followed by acknowledgment and references.

RELATED WORK

Over the past few years, extensive research on the above-while working on the project objectives, a range of literature were reviewed. The authors Diwanji et al. (2018) proposed a chatbot system for flipped classroom settings where the data were collected through many types of research finding their strength. Moreover, they discussed many features that could offer by the chatbot system and how useful is a chatbot system for the modern world, but they lack on providing information on the implementation of the chatbot systems. There is a need to research based on the implementation of the chatbot system instead of researching how chatbot systems would be useful. The authors Hajare et al. (2018) proposed a chatbot system for education that could provide information for the study content asked by the user from an external source or local database, but the popular chatbot systems (Siri, Alexa, Ok Google, etc.) available now are also being used by the students for acquiring information on any type of study content asked by the user and are free of cost.

Although, the proposed chatbot system by Hajare et al. (2018) does not link to any other application in educational fields. And the most popular chatbots (Siri, Alexa, etc.) can deliver information to the user for any study contents from the internet and will invalidate the proposed chatbot system. The authors Naveen et al. (2016) proposed a chatbot system for visually impaired people where the system lacks security for the application and needs an authentication procedure using their mobile number or email address. The research article entitled 'Chatbot based College Information System' by Ram (2019) proposed a chatbot system for the college information system to provide student information by linking with Facebook accounts. However, there is a lack of security as hackers can steal college/institutional data by hacking Facebook accounts hence, a chatbot system without linking social media accounts of the user should be developed.

METHODOLOGY

This research work was supposed to successfully develop a small-scale chatbot system for an educational institute like Middle East College in Oman. The first step of the project was the requirement analysis phase. The product requirements and specification analysis phase were the next step where the complete specification of the project system was analyzed with the help of visuals, block diagrams, schematic diagrams, and flow charts available from the literature. The architecture or high-level design was the next step which defines how the project system should perform the functions and fulfill the design. The design work is divided into various functional sub-blocks to determine the working capability of the project system altogether. The details of the design phase define the delivery of algorithms for each architectural component. Each of the sub-blocks were joined to form a modular design.

The production, operation, and maintenance phase dealt with the enhancements and corrections. The system acceptance testing phase inspects the entire software system and how they are going to perform in the project environment. The integration testing phase checked the modules whether they are interconnected correctly or not. The unit testing phase checked the working of the modules as expected. The coding phase converted the module design into codes and transformed the algorithm, data structures, and programming codes into the required project system. The unit testing, integration testing, system testing, and acceptance testing were considered in the validation phase where the proposed project system was validated through these testing stages.

Our designed system works as TCP server and client connection. The Raspberry Pi 3 (RPi 3) acts as TCP server and the android application acts as TCP client. Users can send the inputs to the RPi 3 using the android user interface application through the same Wi-Fi network. The Raspberry Pi 3 requires a Raspbian Operating System (OS) installed on a Secure Digital (SD) card. RPi 3 boots up Raspbian OS from the SD card and will be opened with desktop idle. It is required to connect to a wi-fi router which provides internet connection and it requires Python idle ver-2.7.16 programming tool for programming the server codes. The android based user interface application can be designed using the App Inventor

Designer tool and App Inventor Block Editor Tool in MIT app inventor 2 online platform. MIT app inventor 2 is a free online platform for building android based user interface applications.

App Inventor Designer tool is used for selecting the components required for the application and App Inventor Block Editor tool is used for assembling the program blocks that specify how the components should work or behave. After designing the application, it is possible to create a stand-alone android application which can be installed on any android phones. By default, it will save the application in .apk format. To initialize the user interface application, it is required to install the designed application on an android device. Also, it is required to connect to the same wi-fi router in which the server is connected. Wi-fi router requires to provide internet connection for the server and the client. The user interface application is required to initiate TCP socket connection by using the server's IP address and port number for the working of chatbot system.

Raspberry Pi 3 used to develop the chatbot system works as Python Transmission Control Protocol (TCP) server. RPi 3 generates a server socket when it starts. Socket is said to be as an end point of two-way communication link between two different programs running on a network and goes on wait mode after initializing the server and client connection. A socket connection will be created by the client using server Internet Protocol (IP) address and port number, it tries to establish a communication link with the server. RPi 3 will wait for a service request from the client. Usually, the waiting of the server is said to be as "listening" for an incoming client. Generally, the TCP works with the IP address which explains how the computers send data packets to each other. The server receives the input data in text format sent by the user and then server processes it. The server sends back the output data in text format to the client. The datasheets which shows the working specifications and functionalities for Raspberry pi 3 are available.

Wi-fi router is used to provide internet connection for TCP server and client to transmit the data between them and the actual data transfer between the server and client is performed using Internet Protocol (IP). It is important to use the client socket extension while designing the chatbot system user interface in MIT app inventor 2. The connection will not be possible if the user interface application is developed using a damaged client socket extension component or without client socket extension. The input parameters for working of the project system are the wi-fi router which provides internet connection for both server and client, android based user interface application and server.
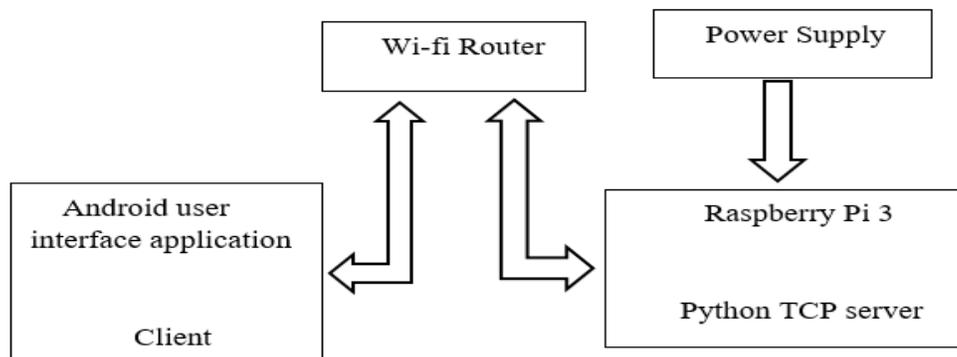

Figure 1: System Block Diagram

The figgure 1 shows the system block diagram. Raspberry Pi 3 works as the TCP server powered by Python which processes the input and provides the output data through Wi-fi. Raspberry Pi 3 and android application are connected to the same Wi-Fi router which provides internet. In our design, the Python TCP server, client, and Wi-fi routers are the three major elements involved in the chatbot system development. The System Raspberry Pi 3 works as a Python TCP server. The android based user interface application works as a client and is developed using MIT app inventor. The Python TCP server is developed using Python idle ver-2.7.16 programming software. Raspberry Pi 3 is powered ON via a 5V power supply. The android based user interface application is used to send input data as voice or text format to the server. The processed data is then sent back to the user as output and is displayed on the user interface application also output is read aloud by Text to speech engine following the flow chart illustrated in fig.2.Our chatbot system follows ALICE algorithm and works using three different pattern matching techniques.

• Keyword Reduction Technique

The chatbot system provides the result according to the keywords present in the query.

• Synonym resolution technique

The chatbot system returns the same output for various queries asked by the user with the similar synonyms.

• Symbolic reduction technique

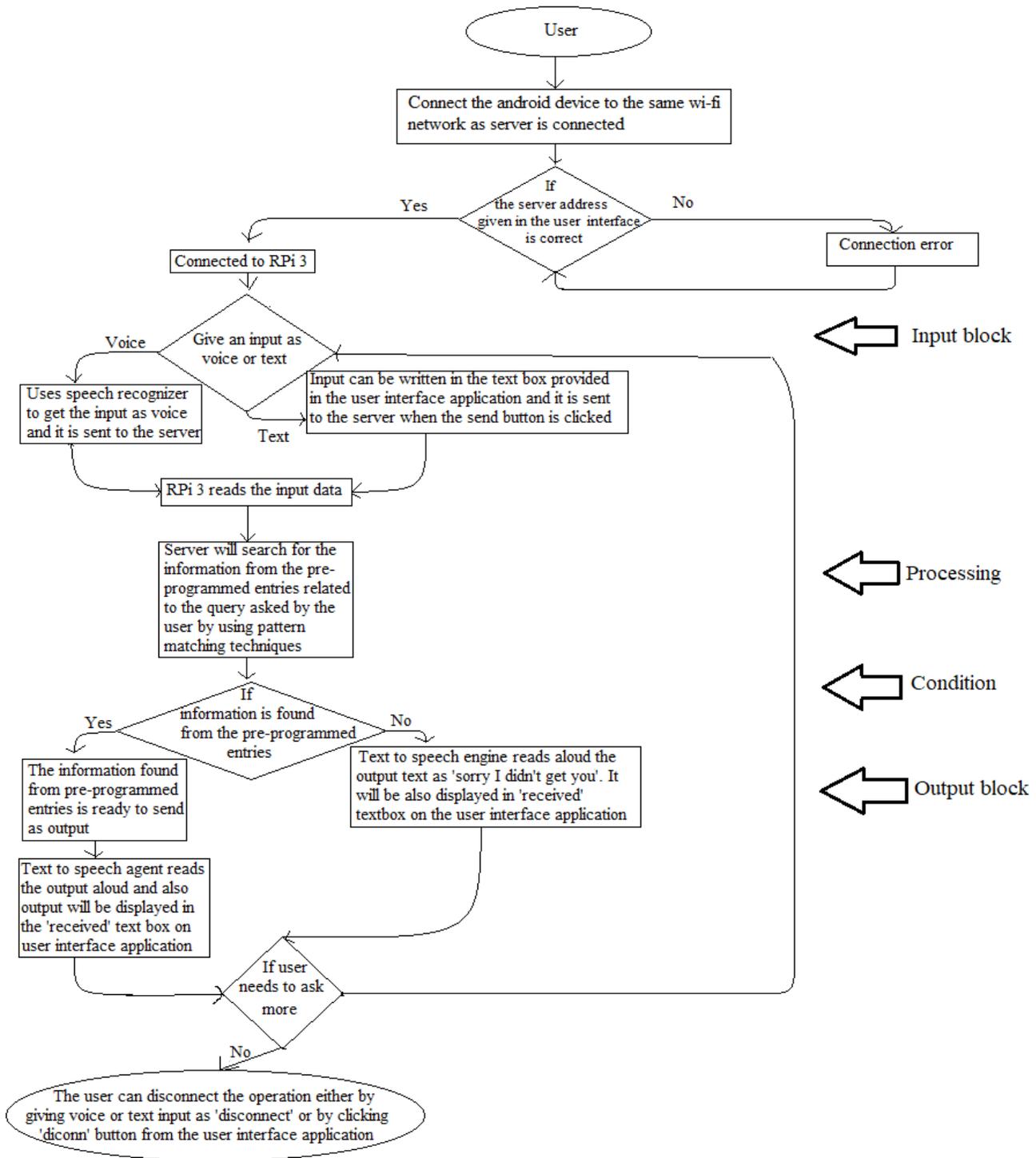This technique is used to reduce the complexity queries asked by the user.

Figure 2: System Flow Chart.

RPi 3 as Python TCP server and the android based user interface application as client, are the key elements for the functioning of the chatbot system. The flow of the process of the chatbot system has been illustrated step by step and described in fig. 2. The process starts by initializing the android based user interface application and connecting it to the same wi-fi network that is connected by RPi 3 (server). The user interface application requires a socket connection to the server address for the working of the chatbot system.

Thus, providing a wrong server address in the user interface application leads to an error socket connection. The user interface application takes the input data as voice or text format from the user and will be sent to the server. This is represented as the input block in fig. 2. Voice input can be given to the chatbot system using the speech recognizer button

from the user interface application and also text input can be given in the normal form of texts. RPi 3 will read the input data and processes it. The collecting of information related to the queries asked by the user from the pre-programmed entries using pattern matching techniques are indicated as processing block in fig. 2. Pre-programmed entries are the local database that contains a set of pre-programmed queries and answers. The condition in which the server decides whether the information related to the queries asked by the user is present in a local database or not is indicated as a condition block in fig. 2.

The output will be obtained from the user interface application invoice format using text to speech engine and in text format as well. This is indicated as the output block in fig. 2. The input block involves a decision box because of selecting either voice format or text format for sending the inputs and the condition block also involves a decision box as the server is deciding whether the information is present in the database or not. The processing block and output block involves a rectangle box, and all the boxes are followed with flow directions.

DISCUSSION OF RESULTS

The user should give the proper server address of the Raspberry Pi 3 in the user interface application to establish a socket connection.The raspberry pi is the main device used for the working of the developed project system which takes the input data from the user through Wi-Fi and provides the result to the user through Wi-Fi itself as illustrated in fig.3.

The user interface application acts as the client and the raspberry pi acts as the Python TCP server. The Raspberry Pi 3 is the key device used for developing the project system and it comes with an on-board 802.11n Wireless Local Area Network (WLAN) adapter, so an extra purchasing of Wi-Fi dongle is not required. Wi-fi dongle is a pocket-sized device that helps the user to access the internet on the go. Both the client and server must be connected to the same wi-fi router and the client requires to connect to the server's IP address. Raspberry Pi can be powered with a micro USB cable by connecting to a 5V AC adapter. High Definition Multimedia Interface (HDMI) cable is used to transmit the video/audio signal to a monitor so as the developer can get the program operating window.

A keyboard and a mouse are connected to the USB terminals. Just top to the micro USB power connector on the raspberry pi circuit board, an SD card slot is available where an SD card with the loaded Raspbian OS can be inserted. Once the SD card is inserted, the Raspberry Pi will boot Raspbian OS from the SD card. The Python 2.7 programming software is used to program the chatbot system. The fig. 3 represents the diagram of system testing.
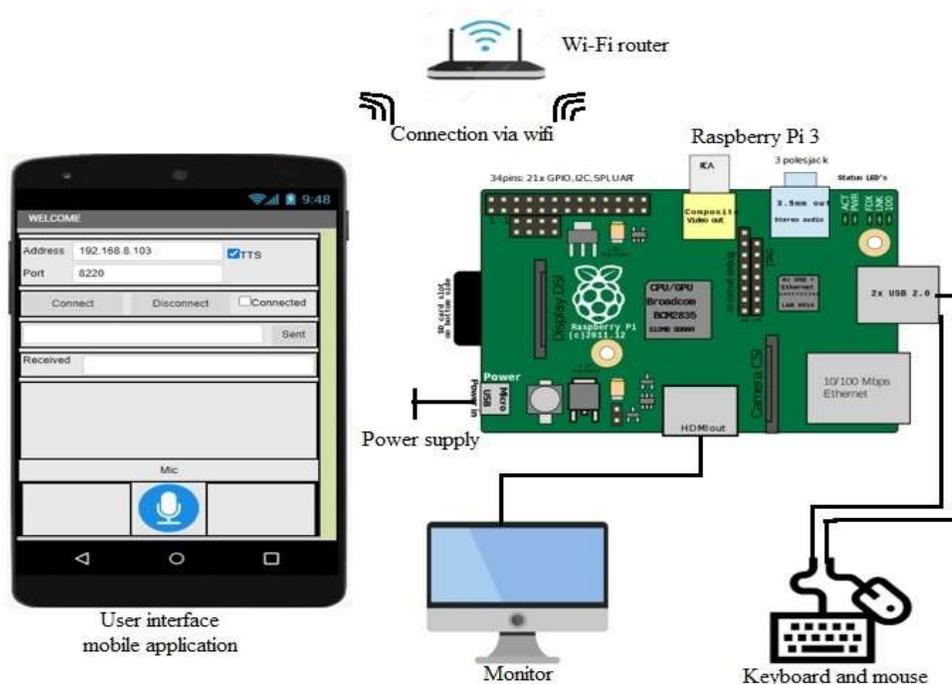

Figure 3. System Testing

Fig 4 represents the app inventor designer window, and it has four types of sections. The palette section contains many tools like sensors (accelerometer, barcode scanner, magnetic field sensor, etc.), connectivity tools (web, Bluetooth server, and client, etc.), etc. The user interface tab from the palette is used to design the screen view of the user interface. By

clicking on the user interface tab in the palette, it will display the elements like a spinner, slider, label, button, etc. So, just dragging them to the viewer section is enough to design the screen view of the user interface.

Figure 4: App inventor designer and blocks editor window

The Components section shows the components used from the palette section for designing the application. The Properties section is used for editing the names, background color, text tool which contains font style, text size, alignment, etc. The speech recognizer and text to speech tools will be in the media tab present in the palette section. These tools are used for the chatbot system to send the input as voice also receiving the output as voice. The text to input (TTS) is also called read-aloud technology which reads the digital text aloud. So, whatever the text obtained at the output will be read aloud using this tool in the developed chatbot system. In android mobile phones by default, speech recognizer and text to speech uses Google TTS engine. Also, in this developed chatbot system, the Google TTS engine is used for sending and receiving voice inputs.
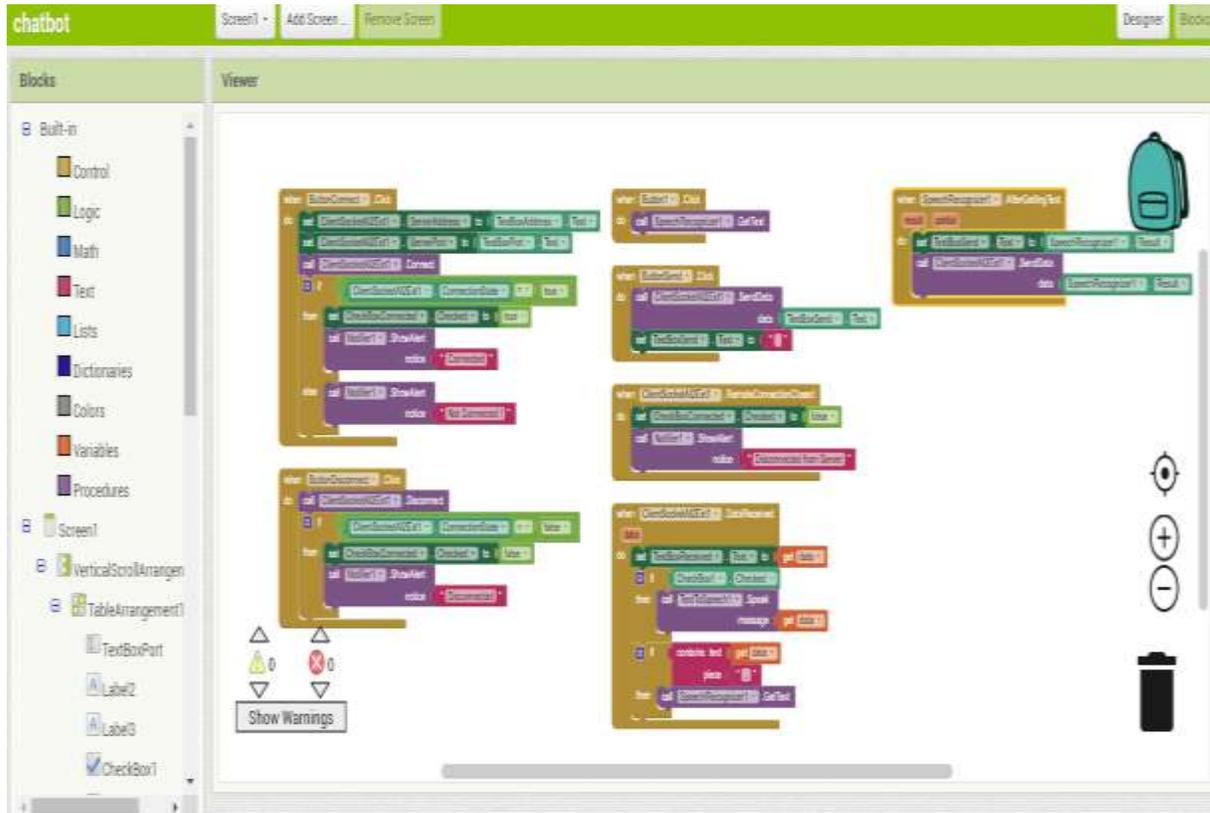

Fig. 5: App inventor designer and blocks editor window

The client socket extension is a key component used for server-client communication. So, a socket extension is used for the connection between the user interface which acts as the client and the Raspberry Pi 3 which acts as a server. The

server establishes an address (binds) first, which the client can use to find the server. The server waits for the client to request a service. Therefore, the socket is the component responsible for the connection between client-server hence the client-server data exchange takes place. It is required to provide the server's IP address on the server address tab in the client socket extension tool menu before building the android user interface application. Then only, the client can connect to the server through the socket.

The fig 4 and 5 represents the blocks editor window. Using blocks editor window, the working of the system can be programmed or assembled using some fitting pieces structures like puzzles instead of programming codes. The results of various test performed are provided in table 1.

Table 1: Tests and Results

Test cases	Queries asked by the user	Desired response from the chatbot database	The response given by the chatbot system
T1	Which semester is this now?	'This is spring 2020' or 'spring 2020'	This is spring 2020
T2	What is my attendance for vlsi module	'Your attendance is 75 percent' or '75percent'	Your attendance is 75 percent
T3	Again vlsi attendance how much?	'Your attendance is 75 percent' or '75percent'	75percent
T4	Hi can I have an appointment for Asif sir	'appointment booked for coming Tuesday 9am to 10am' or 'booked for coming Monday from 3pm to 4pm' or 'no problem appointment booked for coming Thursday 11 am to 12 pm'	no problem appointment booked for coming Thursday 11 am to 12 pm
T5	When is the submission deadline for radar assignment?	'the submission deadline for first assignment will be on 12 June 2020 and second will be declared later' or 'the first assignment submission deadline will be on 12 June 2020 and second will be declared later'	the first assignment submission deadline will be on 12 June 2020 and second will be declared later
T6	disconnect	'disconnected'	disconnected
T7	hehe	'sorry I didn't get you'	Sorry I didn't get you
T8	Any activities going in IBR building	'A workshop based on how to prepare a good resume is going on in room number 213' or 'you do not have any activities currently now but a business quiz will be conducted at 3 o clock by business club' or 'a quiz will be conducted by electronics today at 4 clock'	a quiz will be conducted by electronics today at 4 clock
T9	Is there any program going on in ibr building	'A workshop based on how to prepare a good resume is going on in room number 213' or 'you do not have any activities currently now but a business quiz will be conducted at 3 o clock by business club' or 'a quiz will be conducted by electronics today at 4 clock'	A workshop based on how to prepare a good resume is going on in room number 213
T10	Can you please check if Muhammad Bashir sir is available this and please book an appointment for him	'appointment booked for coming Tuesday 9am to 10am' or 'booked for coming Monday from 3pm to 4pm' or 'no problem appointment booked for coming Thursday 11am to 12pm'	booked for coming Monday from 3pm to 4pm

In our designed system, the server and client must be connected with the same Wi-Fi network for enabling a TCP connection. The user interface application only works in Android OS. The RPi 3 board becomes hot quickly while working with it, so a cooling fan is required to be kept along with the RPi 3 board. The user interface application would not connect to the server's address of RPi 3 when a wrong IP address of the server is provided in the user interface application, hence it shows socket error connection. By giving input 'disconnect' as text or voice will disconnect the user interface application with the RPi3. The students can interact with the chatbot system using android based user interface application by asking queries related to student information and learning activities.

Hence, it enhances the student's engagement in all learning activities and improvising the learning process. Further development in the programming of the chatbot system and in the database will make the chatbot system more flexible.Further gathering of metrics, collecting surveys, and continuous monitoring of the success of the chatbot system can improve itself over time. The application of the developed chatbot system in this project is for education. The developed chatbot system can be also used for home automation, conversational commerce, customer service, personal assistance, etc. by using appropriate programming.

CONCLUSIONS

The developed chatbot system was aimed at enhancing the student's engagement in all educational activities. A range of literature review information is obtained from detailed research based on the topic 'AI based chatbot systems in the educational field'. The developed chatbot system follows ALICE program, and uses three different types of pattern matching techniques which are keyword detection technique, symbolic reduction technique, and synonyms resolution technique. The system working is based on Transmission Control Protocol (TCP) server to client connection. Android-based user interface application was developed using an online platform which is MIT app Inventor 2. The TCP server is developed using Python ver-2.7.16 software on Raspbian OS. The server to client connection is established using a socket. In a designed system, both the functional units are required to connect with the same network which enables the TCP/IP data transmission over the same network. The results are obtained from the android based user interface application and using Python Interpreter. Python Interpreter lists out the queries asked by the user and responses by the chatbot system for every chat session.

The results obtained from the user interface application and python interpreter are verified with the three different types of techniques used for developing this chatbot system. The keyword detection technique is used to detect the main keywords present in the sentence asked by the user and providing a suitable response from the template which is already pre-programmed in the programming file. The symbolic reduction technique is used to reduce the complexity of the sentence by splitting it into individual words. Thus, keywords can be detected and provide the output from the related keyword's template. The synonyms resolution technique is used to find the keywords with the same meaning. Hence selecting it as a keyword and finding the information from the related keyword's template. The results are verified with these techniques using the developed chatbot system by asking different queries.

The project can be further improved by extending the programming script of the chatbot's server and the pre-programmed entries. The developed project system can be implemented in educational institutions to enhance the student's engagement in all learning activities. The institution which is implementing this chatbot system can use their local database rather than pre-programmed entries, thus the chatbot system can become more efficient in providing the output. Designing user's authentication will improve the privacy and security of the users. Further development in Natural Language Processing (NLP) will make the chatbot system more efficient, user friendly, and more flexible by feeding them with new information over time.

REFERENCES

Diwanji, P. H. (2018). [online] available from https://pdfs.sema. *'Enhance Classroom Preparation for Flipped Classroom using AI and Analytics'. Proceedings of the 20th International Conference on Enterprise Information Systems.*

Fawaz, A., Ali, M. B., Adan, M., Mujtaba, M., & Wali, A. (2021). A Deep Learning Framework for Efficient High-Fidelity Speech Synthesis: StyleTTS. *iKSP Journal of Computer Science and Engineering*, *1*(1).

Hajare, A. B. (2018). Chatbot for Education System. *International Journal of Emerging Technology and Computer Science , 3*(2), 43-48.

Lerma, J. M. C. (2021). SmartHipona: Distributed Computer System to Record the Academic Performance in Real Time. *iKSP Journal of Computer Science and Engineering*, *1*(1).

Naveen, K. L. (2016). Android Based Educational Chatbot for Visually Impaired People. *International Conference on Computational Intelligence and Computing Research* .

Oh, K. L. (2017). A Chatbot for Psychiatric Counseling in Mental Healthcare Service Based on Emotional Dialogue Analysis and Sentence Generation. *Oh, K., Lee, D., Ko, B., and Choi, H. (2017) 'A Chatbot for Psychiatric*

Counseling in Mental Healthcare Service Based on Emoti IEEE International Conference on Mobile Data Management (MDM) 1, 371-37.

Ram, M. (2019). Chatbot based College Information System. *International Journal of Multidisciplinary , 4*(3), 109-112.

Sarthak V., D. S. (2017). Artificial Intelligence Chatbot in Android System using Open Source Program-O. *(2017) 'Artificial Intelligence Chatbot in AnInternational Journal of Advanced Research in Computer and Communication Engineering , 6*(4), 816-821.

Sonali, M. a. (2010). Advancements in the V-Model. *International Journal of Computer Applications, 1*(12), 29-34.

Tim, H. (2019). *Can a computer fool you into thinking it is human?* (BBC News) Retrieved 1 14, 2020, from https://www.bbc.com/news/business-49344596