

A QNA Bot Using NLP

Prof. T.P. Vaidya

Computer Engineering Sinhgad College of Engineering Pune, India tpvaidya.scoe@sinhgad.edu

Pradip Kokate

Computer Engineering Sinhgad College of Engineering Pune, India pradipkokate15756@gmail.com

Shalaka Kudale

Computer Engineering Sinhgad College of Engineering Pune, India shalaka.12kudale@gmail.com

Akash Jadhav

Computer Engineering Sinhgad College of Engineering Pune, India

akashjadhav004@gmail.com

Aaditya Kulkarni

Computer Engineering Sinhgad College of Engineering Pune, India

kaaditya598@gmail.com

Article Info

Volume 83

Page Number: 2233 - 2238

Publication Issue:

March - April 2020

Article History

Article Received: 24 July 2019

Revised: 12 September 2019

Accepted: 15 February 2020

Publication: 18 March 2020

Abstract

The interview process is a time-consuming process; therefore, many companies prefer to visit those campuses where it can get a greater number of freshers. Even though some colleges have eligible candidates, the company doesn't prefer to visit it as such candidates are few and company can't afford to spend their valuable time for a few numbers of students. In such scenario an AI bot can be developed that automates the interview process. The company doesn't have to visit the college, just by using the bot it can get the desired candidates from anywhere. It would also create an opportunity for the candidates who wish to work in such companies. Thus, creating a win-win situation for both. Such bots could also be used by colleges to train their students for the placement activities.

Keywords – *Natural Language Processing, Artificial Intelligence, QNA BOT, Natural Language Toolkit, CNN.*

I. INTRODUCTION

In an increasingly complex world, automating the interview process is necessary for the benefit of students. Most of the interview taken by companies requires a lot of time and manpower. Such interviews may even be biased. The interviewer may select someone he likes or wouldn't select someone if he dislikes irrespective of the eligibility of the candidate. The core problem is to solve problem of student living in a rural area or the areas where the best companies can't go just for the interview process. Our application act as a bridge between those students and companies, it will also help people to improve their communication skills along with technical skills. It is easily accessible to everyone. Our application includes chat bot with Facial

Recognition, Body Gesture detection which will be helpful to overcome the problems.

A chatbot is used for the conversation between humans and machines via auditory or textual methods. In our application the conversation will take place between an Interviewee and Interviewer. There are many benefits of using a chatbot, it works 24/7, it never judge a person based on caste, race and ethnicity. Chatbots are a quick way to inform the candidates about their selection or the next steps in the process.

It involves building a single platform which will work as an assistant in improving skills of the students. There will be a web-based platform which will be available to both students and company. The application will be different login for different companies for the purpose of security. Even different test for practice will be available for the students and result will be displayed on the

screen, the record will be kept so as to analyse the question and answering will be done with the help of Speech to Text and Text to Speech i.e. by using NLP (Natural language processing) so as to fell the real scenario of interview. Voice recognition and emotion detection will be done with the help of NLP libraries.

Natural language processing (NLP) is a technique used by which the system can understand the instructions for control the text or speech. A Text-to-speech system is developed, it converts text into spoken words, by processing and analysing it using Natural Language Processing (NLP) and Digital Signal Processing (DSP) technology, it helps to convert the processed text into speech representation of the text. Here we have developed a useful text-to-speech synthesizer in the form of a simple application that converts input text into synthesized speech and reads out to the user which can then be saved as an mp3 file and the saved mp3 file is used during the interview to recognize the user with the help comparison between his/her current voice and the saved voice.

Speech Recognition is the simple method in which the spoken language or the audio file is converted into text/words by the computer. This is also known as computer speech recognition, automatic speech recognition, or speech to text.

II. LITERATURE SURVEY

A. Available Technologies:

1. Mock AI – Mock AI is a tool used by amcat. Used to help to fight any second-thoughts which come before the interview. It is a master product, that has helped thousands of students for their interview preparation. Mock AI tests student's interview skills on 20 sought job roles in country. The amcat mock-AI creates a stress-free and almost real-time interview environment. The software uses Artificial Intelligence and Machine Learning to provide feedback regarding facial expressions, body gesture and voice modulation.

The feedback helps to get insights into strong and weak areas.

2. Allyo – It is used for automating interviews via SMS, used to find more qualified candidates. Has over 10 million conversations with candidates till now and boost different customers like FedEx and Cheesecake Factory. It also allows for surfacing the previous candidates.

B. Related Papers:

1] Summarization of spontaneous speech:

Speech summarization extracts important information and removes redundant and incorrect information, necessary for transcribing spontaneous speech. Efficient speech summarization saves time for reviewing speech documents and improves efficiency of document retrieval. Speech summarization is done on the basis of unit extraction and concatenation [5].

Methods for presenting summaries by speech can be classified into two categories: 1) presenting simply concatenated speech segments that are extracted from original speech or 2) synthesizing summarization text by using speech synthesizer. Major problem of extracted speech segment is how to avoid unnatural noisy sound caused by concatenation [5].

Speech to text summarization includes two-stage automatic speech summarization method consisting of important sentence extraction and word based sentence compaction. For the case of presenting summaries by speech, the speech to speech summarisation works on 3 kinds of units - sentences, units, words and between filler units - were investigated as units to be extracted from original speech and concatenated to produce summaries [5].

2] Speech recognition using deep neural network:

Speech recognition system was introduced at Bell Laboratories which could recognize single voice speaking digits. After Ten years, IBM introduced "Shoebox" which used to respond and also understood up to 16 words in English.

Speech in particular being the main method of communication among human beings, received much interest for past 5 decades right from introduction of artificial intelligence. Therefore it is only natural that one of early application of deep learning for speech related application specifically speech recognition. The conventional speech recognition system are based on representing speech signals using Gaussian Mixture Models (GMMs) that are based on Hidden Markov Models (HMMs). There are many examples that have proven that deep neural network yield better results than classical models. In 2012, Microsoft released newest version of Microsoft Audio Video Indexing Service (MAVIS), speech system based on deep learning [4].

Automated speech recognition is capability of machine or computer to recognize the content of words and phrases in uttered language and transform them to a machine understandable format.

III. PROPOSED SYSTEM

A. System Architecture:

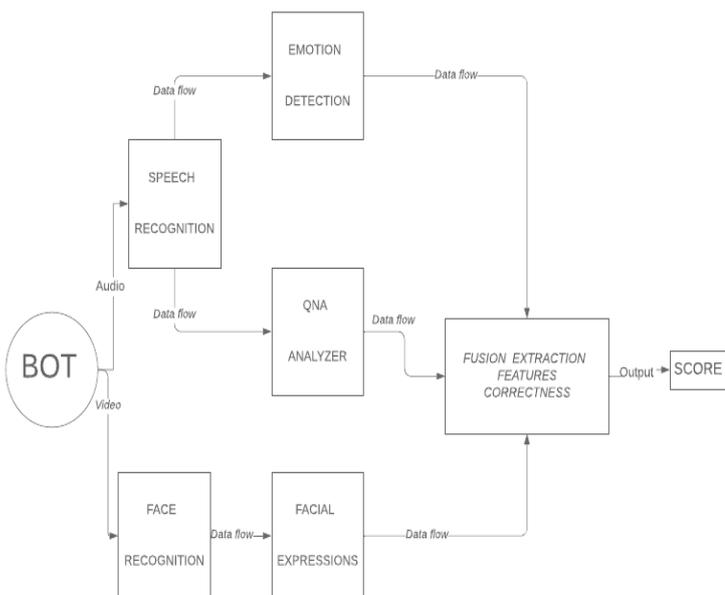


Figure 1: System Architecture of BOT
Proposed System consists of Three modules viz.

1. Voice Detection
2. QNA Analyzer
3. Face Recognition

Voice Detection:

Initially the speech of the user will be recorded and stored; the frequency of the voice will be calculated by using MFCC. The input will be taken as an audio file or speech of the user. During the interview process, the frequency of ongoing speech will be calculated and average of the whole frequencies will be compared with the previous frequency, if both the frequencies are near to each other, it means the user was constant. Along with the frequency calculation the emotion of the user can be detected by converting speech to text and from the text emotion can be detected [1]. By knowing both the outputs i.e. Frequency output and emotions we can judge whether the interviewer confident or nervous. It will judge the student on basis of her/his non - technical skills.

QNA Analyzer:

Questions for the interviewee will be extracted from the firebase database. Question along with the answer is taken and the answer given by the user is compared with the original answer. If both the answers are right then interviewee is scored with plus one else minus one. The output of the QNA analyzer will be total questions asked and questions answered correctly. QNA analyzer will judge the student on basis of her/his technical skills.

Automated Facial Expression:

The process proposes a model which performs analysis of both speech and facial expression to detect the emotion. But the process need to modify motions related to interview process, only identify general actions. It does not take body language into account.

The model used here facial landmarks for facial expression detection. SVM algorithm is used for classification process. The results are expressed as a percentage of all the possible emotions, such as sad, happy, fear, surprise, anger and so on. The results of the experiment confirm

that face and speech emotion recognition was conducted successfully using a smartphone. It was correct in 97.26% of instances when used with standard corpora. However, this is not taking body language into account which is also an important aspect to understand a person's behaviour.

The following are the main features that are included in the application:

1. Cross platform support: Offers operating support for most of the known and commercial operating systems.
2. User account: The system allows the user to create their accounts in the system and provide features of updating and viewing profiles.
3. QA Bot: Question an answer bot along with text to speech and speech to text.
4. Portfolio Management: Once the user has been allocated with different for result of different tests, this module then helps in managing their result to improve him/her.
5. Payment Gateway: This module is free of cost.

B. Methodology

The System consists of methods such as: Speech Recognition and Tone Analyser. The input to the system is the audio taken using the microphone respectively.

The answer given by the candidate along with the tone of answering is analysed. The speech is converted into text and text is compared with the standard answer. If the answer is matched then marks according to the weight associated with it is allocated. Further, based on the tonality of answering the marks are allocated too.

All these marks are combined and total percentage of the candidates according to the biases is calculated. Those who score more than the expected percentage clear the interview.

Interviewers may need a candidate whose technical knowledge is sound or they may even need a candidate whose communications skills are great. There may be different requirements for different companies. So this system also provides

an option to set biases and select candidate according to the biases.

Once the interview is done, the candidates may check their performance; they may look into the details and understand in which part they are strong and where they are lagging. And accordingly, they may improve themselves. This could be an excellent bot for the college students who are willing to improve themselves before giving an interview. This works just like a mock interview.

The output of the Bot will be represented in the form of pie chart, which will define students' performance in each field on the basis of percentage assigned to each field. Result of the user is managed in such a way to maximize his/her profits and reduce risks.

There are multiple API's that can be used in python to convert text to speech, the most common used is Google Text to Speech API or gTTS. In our system we have used pyttsx3 python library to convert text to speech, it works offline and is very easy to use. It supports 2 voices: Male and Female. [8]

Emotion Analysis is done with the help of Parallel Dots API. [6] The input voice is converted into text and on the basis of the text total six emotions are detected. The six emotions are as follows: Angry, Happy, Excited, Sad, Fear and Bored. [7] The video goes to the facial expression and the body language analysis module, where score is given to the candidate according to his/her expressions and body language.

The origin of this product was observed after knowing that the colleges located at rural areas or the companies can't waste their time in travelling to the college, such companies can have access to the students by conducting online interviews. The development of this product has become a necessity for the betterment of the students. The main perspective of this product was to make access to the students to the companies and also

communicate the companies to students. This application not only focuses on the technical skills but also concentrate on confidence of the student.

IV. CONCLUSION

It became difficult to the companies to travel to different colleges and take the interview of

students and select best of them, so that it will be beneficial for the company. But the students in rural areas don't get such facilities or opportunity to become a part of some companies. So to reduce the efforts and time of company and to select the deserving candidate from rural areas is the main aim of our product.

```
C:\Users\Pradip\PycharmProjects\BEP4\venv\Scripts\python.exe C:/Users/Pradip/PycharmProjects/BEP4/python_templates/speech_voice.py
Tell me about OOP concepts
Say Answer...
Now say
Expected Frequency: 46.85625

Yours voice Frequency: 22.325625

Your voice is regular.

Emotions from your voice seems to be like as following:
{'emotion': {'Excited': 0.1259794439, 'Angry': 0.3479545631, 'Sad': 0.1651695593, 'Happy': 0.0698302814, 'Bored': 0.1134961031, 'Fear': 0.1775700492}}

Your answer is: inheritance polymorphism and exception these are the main concepts of object oriented programming language
Standard Answer words(Your answer should contains these words) : ['Polymorphism', 'Inheritance', 'Abstraction', 'Encapsulation']

You have got 50.0% marks for this question

Process finished with exit code 0
```

Figure 2: Output of Speech Recognition

Fig. 2 shows the output of speech recognition module. The input for the speech recognition module will be voice, from the voice it will judge the confidence level of the student. The output

will be both technical score i.e. score of correct answers and non – technical score i.e. it will check whether the candidate is nervous or confident.



Figure 3: Output of Face Recognition

Fig. 3 shows the output of face recognition module. The input for the face recognition module will be video, from the video it will detect facial expressions and capture the images will also help to check whether the student is cheating in the interview or not.

This product helps in providing more employment to deserving students from rural areas. It also helps them to increase confidence level with the help of practice test provided.

V. ACKNOWLEDGMENT

The least we can do is express our gratitude to the following individuals. Their contribution has been noteworthy in the completion of the project "A QNA bot for using NLP". It gives us immense pleasure to acknowledge the innumerable guidance extended by Prof. T.P. Vaidya. She has always enlightened us with her valuable suggestions and ideas. In this endeavor we acknowledge the contribution of Prof. M.P. Wankhede, HOD, Computer Engineering Dept., Sinhgad College of Engineering for encouraging us to perform this research. All his suggestions and advice have always been of great help. We are also grateful to Prof. S.D. Lokhande, Principal, Sinhgad College of Engineering, for his continued and unending

support. We are thankful to Sinhgad College of Engineering, Pune for providing us all the necessary resources. We extend our sincere thanks to the Savitribai Phule Pune University, Pune for granting us permission to work on this research along with our academics. We would also like to thank our family and friends, and all those who helped us complete this research work.

VI. REFERENCES

[1] Partha Mukherjee, Ananya Paul, Soumen Santra, Pubali Chatterjee, Subhajit Bhowmick, Arpan Deyasi, "Development of GUI Text-to-Speech Recognition using Natural Language Processing", IEEE Paper, 2018.

- [2] Esther Ramdinmawii, Abhijit Mohanta and Vinay Kumar Mittal, "Emotion Recognition from Speech Signal", IEEE paper 2017
- [3] Qingli Zhang, Ning An, Kunxia Wang, Fuji Ren and Lian Li, "Speech Emotion Recognition using Combination of Features", 2013
- [4] Ali Bou Nassif, Ismail Shahin, Imtinan Attili, Mohammad Azzeh and Khaled Shaalan, "Speech Recognition Using Deep Neural Networks:A Systematic Review", IEEE paper, 2019
- [5] S.Furui, T.Kikuchi, Y.Shinnaka, C.Hori , "Speech-to-text and speech-to-speech summarization of spontaneous speech", IEEE 2004
- [6] https://apis.paralleldots.com/text_docs/index.html#emotion
- [7] <https://www.paralleldots.com/emotion-analysis>
- [8] <https://www.geeksforgeeks.org/python-text-to-speech-by-using-pyttsx3/>