



Interview Analyzer using Natural Language Processing

Ms. Shrushti Asutkar¹, Mr. Narendra Gawai²

^{1,2}UMIT, SNTD Womens University

ABSTRACT—

A pandemic can change everything and can change the flow of daily work life. And as technological students, we need to be readily available with all the upgradeable methods for smooth flow. An academic or business paper is enhanced by qualitative analysis, rather than quantitative analysis, which is certainly important. As well as adding perspective, adding context, and encouraging thoughtful discussion, it can make a paper more readable and better understandable. As For less measurable data, interviews have proven to be essential. In addition to providing relatable stories and perspectives, they also allow researchers to quote important contributors directly. In interviews, authors can gather a lot of qualitative data that allows them to avoid embellishment and preserve the integrity of their content at the same time. Qualitative analysis is best performed using interview transcripts. A study can be brought to life through interviews with the following sources: Science and inven-tion, Leaders in business, Other study and journal authors, Study or trial participants. This can be obtained by Gathering audio data input through recordings of the interview session which will be stored for further encoding and process, encoding the gathered audio data to mathematical vectors which can be used for further assistance in generation of text then we convert encoded audio to text and create vectors from given text .

Index Terms—Natural language processing, Similarity cosine.

1. INTRODUCTION

TO assess qualitative interview data, interview transcripts are frequently employed. For transcripts to be obtained, they must either be well-written or recorded and then tran-scribed, respectively. An interview is transcribed by playing the audio again and typing in every word, noting who spoke which words, to provide a comprehensive written copy of the recorded interview. Aiming for a verbatim transcription, or one that accurately captures every word stated during the audio recording of the interview, is generally the best course of action. The textual transcription of an interview should, if at all possible, also capture nonverbal replies. We can observe the gestures, the voice tones, and any possible accentuation used by responses. Upon the completion of the interviews, the next step is to store and analyze what people have told the interviewer.

There are specific criteria on how to analyse interviews that should be followed regardless of the project, whether it be a simple or difficult one. Also, a thorough analysis makes the findings more trustworthy by allowing others to see how the interviewer came to their conclusions about their participants. All interviewers can be given a questionnaire, which will help with further analysis and prevent any partial treatment of the interviewee. It will also enable the interviewer to draw conclusions about the interviewee's knowledge by comparing the similarity between the questionnaire and the provided answers.

By offering a common questionnaire and quick findings based on similarity percentage, this paper provides a tool that would allow us to conduct fair interviews and derive solid conclu-sions.

A. Objective

This paper focuses on sytem which aims to analyze an-swears used for measuring different aspects of knowing of an individual . Enables the modern recruiter (HRs) to find and engage with qualified candidates fast and the companies to hire the suitable candidate without being biased followed by an increase in engagement with qualified candidates through intelligent guidance to design and automate. Fair judgment and assistance in understanding the depth of knowledge of the candidate is the prime target

B. Scope

- To ask each individual the same job-relevant questions. Streamlining the process and minimizing bias.
- To ask the questions the interviewer wants answers to which enables them to evaluate the interviewee for the precise skills which are expected.
- Transcripts and reports are easily shareable and effort-lessly exchange information to get the best out of the data.
- Virtual interviews can be conducted anywhere with the ultimate convenience for the interviewee, especially dur-ing multi-phase hiring processes.

II. PROBLEM STATEMENT

One of the most widely utilized methods for choosing employees is the interview. A structured interview is one in which a candidate is put through a predetermined list of questions in a specific order. Structured interviews are typically better predictors of which candidates will make good employees. Interviews can range from completely unstructured and free-form conversations to structured ones. In order to gather information on a topic, a structured interview focuses on asking questions in a predetermined sequence. The interview is one of four varieties. Candidates frequently get the same questions in job interviews because they might sometimes take a formulaic approach. This is particularly true of seasoned recruiters, who frequently feel certain enough during their interview skills to simply wing the interview, and consequently, use the same tried and tested questions time and again but this is not the right approach. Every interview should be customized for the role, and you should prepare a list of questions that

are relevant to that position only. The bare minimum is a precise set of targeted queries. The responses to these specific questions don't alter significantly, even if the interview panel does. Nonetheless, the interview panel could comment on a candidate's response that, had the panel been different, might have been better or worse. A more effective method of interview evaluation may be useful to standardize the hiring procedure on a broad scale. Maintaining the same interview panel for several applicants who are competing for the same position will be a very taxing and time-consuming operation for the panel and it will be difficult to gather people to sign up for the process.

III. RELATED WORK

Atlas.ti is a software program used to analyze and manage data from different sources, including interview transcripts. It offers powerful features to help researchers identify patterns, themes, and relationships in their data. The first step is to import interview transcripts into the software and then code the data by identifying segments related to a particular theme or topic. Atlas.ti allows the user to organize codes and build a comprehensive codebook, search and filter data based on specific codes or themes, and use visualization tools such as network views, word clouds, and matrices to identify relationships between codes and themes. Finally, researchers can generate reports and export data in various formats, making it easy to share their findings with colleagues or incorporate them into research papers.

NVivo is a popular software tool for qualitative data analysis that allows researchers to organize, code, and analyze data from various sources such as interviews, surveys, and social media.

MAXQDA is a software tool that helps researchers analyze qualitative data. It includes features like coding, data visualization, and text analysis.

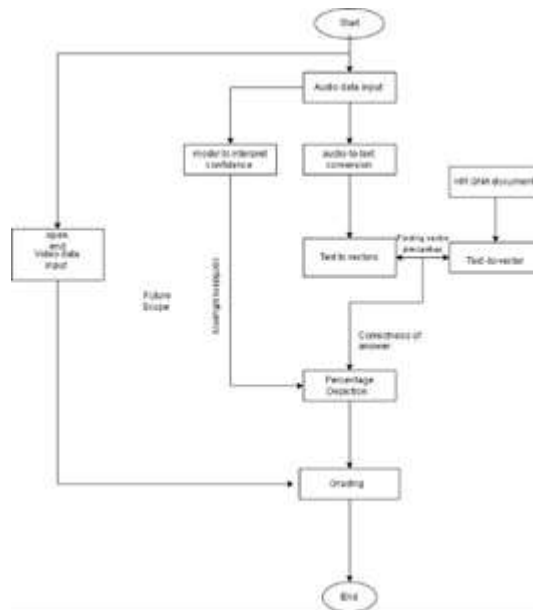
A. Overview of existing system of the interviewee . This audio data needs to be encoded to mathematical vectors via Speech2TextProcessor module of transformers library which can be used for further assistance in generation of text . Text is generated from the encoded audio , and vectors are created from the generated text. In order to conduct equal interviews for all a common questionnaire shall be provided .The analyzer will go through the PDF file that HRs have provided and generate a dictionary of questions and answers. Next, it will create vectors of the question and answer text to compare with the candidate's response. The similarity shall be calculated with the help of cosine similarity where cosine is the the angle between two vectors. The greater the value of θ , the less the value of $\cos \theta$, thus the less the similarity between two documents. The candidate receives a grade based on $(\cosine\ distance) \times 100$ of how similar their response is to the intended response.

The current system of interview analysis has its own final analysis and data-based conclusion. The current approach requires an interviewer to physically interview the interviewee, who is evaluated based on how well he or she performed during the interview process. Conducting physical interviews can be time-consuming due to travel, and situations may arise where an eligible candidate is not available in the same city.

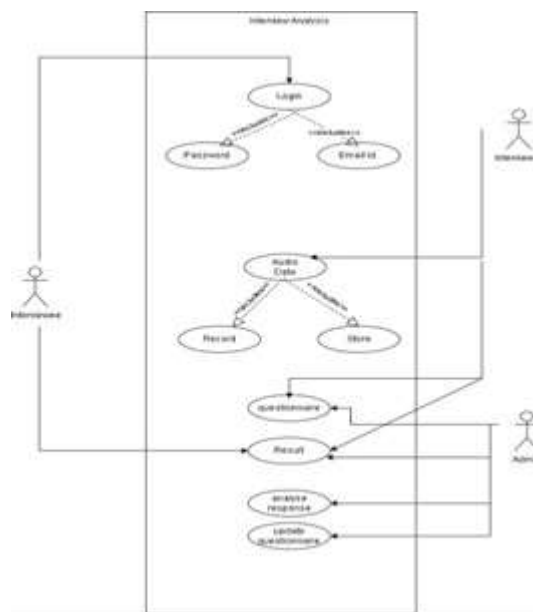
With the occurrence of the pandemic, many tasks had to shift to an online platform, leading to a new approach to making everything online possible. Manual interviews can occasionally fail to provide accurate results because the level of analysis can vary depending on the interviewer and the questions, diminishing the accuracy level.

B. Proposed system

This project aims in analyzing and developing a model which helps with unbiased interview conclusions and faster delivery of results. For the same, the process begins with collecting audio data through interviews with the consent of the interviewee . This audio data needs to be encoded to mathematical vectors via Speech2TextProcessor module of transformers library which can be used for further assistance in generation of text . Text is generated from the encoded audio, and vectors are created from the generated text. In order to conduct equal interviews for all a common questionnaire shall be provided .The analyzer will go through the PDF file that HRs have provided and generate a dictionary of questions and answers. Next, it will create vectors of the question and answer text to compare with the candidate's response. The similarity shall be calculated with the help of cosine similarity where cosine is the the angle between two vectors. The greater the value of θ , the less the value of $\cos \theta$, thus the less the similarity between two documents. The candidate receives a grade based on $(\cosine\ distance) \times 100$ of how similar their response is to the intended response.



Proposed Architecture (2)



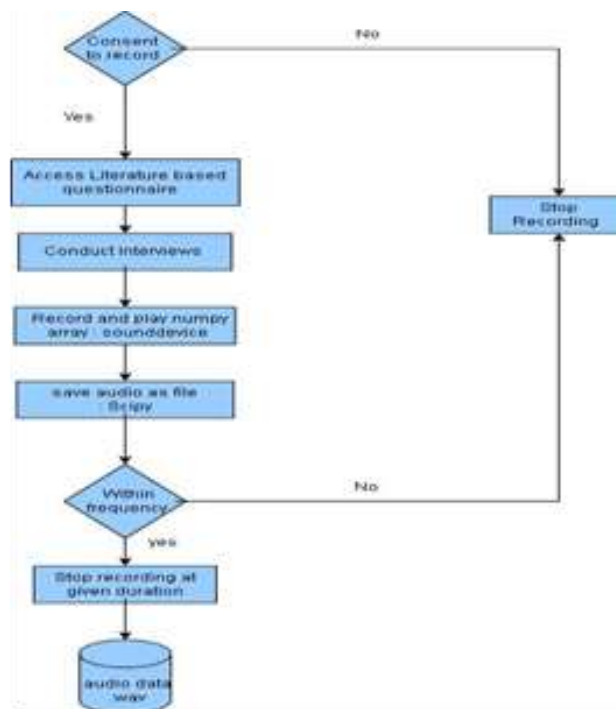
Use case (4)

IV. METHODOLOGY

The paper highlights the process of creating an unbiased interview model that delivers results quickly. Audio data is collected through interviews, encoded into mathematical vectors, and used to generate text. A common questionnaire is provided for equal interviews, and the analyzer creates vectors of the question and answers text to compare with the candidate’s response using cosine similarity. The candidate’s response is graded based on how similar it is to the intended response.

A. Audio Data Storing

We utilize the “sounddevice” module, which has functions to record and play numpy arrays, to capture the audio data obtained from the interviews and save it in files for later analysis. We determine the frequency and duration. The function to a numpy array. In the end, sounddevice.wait() is used to wait until the recording is complete and for saving the recording as the wav file we write the write() function. rec(int(duration*fps), samplerate = fps, channels = 2) is used to deliver audio data

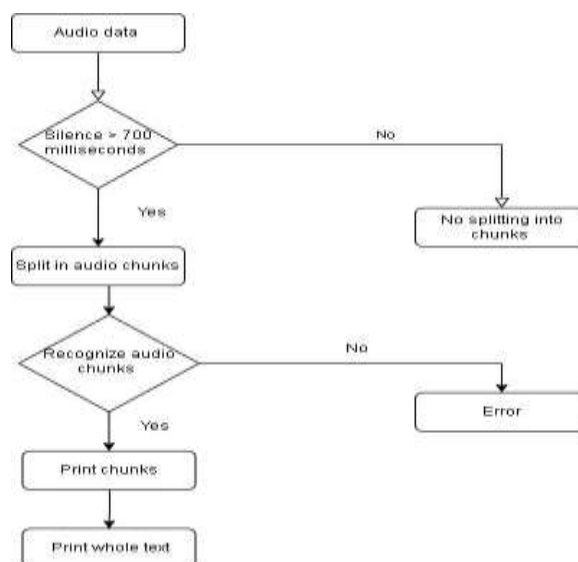


Flowchart : Audio Data Storing

B. Converting audio data to text

The length of the audio file used as input reduces speech recognition’s accuracy. As a result, we must split the audio file into smaller pieces before identifying those pieces. The result of each of these pieces can then be concatenated and printed. By breaking up the audio recording into segments of constant size, we run the risk of losing some important words and possibly breaking up phrases. Unfinished sentences cannot be used because the audio recording could end before they are finished. The audio file can be separated based on quiet as well. Between each uttered word, there is a little silence.

Based on these pauses, we may divide the audio stream into parts, which will allow us to treat each segment separately and combine them to create the desired output. This method is more accurate than the previous one since the audio chunk will contain the complete line without any breaks. In this manner, there is no need to divide it into units of fixed length



(7)

Flowchart : Audio Data to Text (8)

C. Create vectors from given text

The Doc2Vec model is similar to Word2Vec, both use unsupervised learning to extract word embeddings from a corpus. However, Doc2Vec uses a technique to extract paragraph vectors rather than vector representations of the entire corpus. It does this by using the paragraph id as a distinctive vector, making it different from Word2Vec.

D. Create vectors from the HR doc

Analyzer will read the pdf file uploaded by HRs and create a dictionary of questions and answers, now create vectors of the question and answer text to compare with the answer given by the candidate.

E. Finding similarity of answers

We employ the cosine similarity metric, which can be used to determine how similar two data objects are, to compare candidate responses to real replies, independent of size. Python's Cosine Similarity function can be used to compare two texts for resemblance. Every data object in a dataset is treated as a vector via cosine similarity.

V. IMPLEMENTATION

Development can be divided into three parts

1. Data Collection
2. Vector formation
3. Score calculation

In Data Collection, we record and save input audio file collected via a microphone. This is divided depending on silence. People pause briefly between statements in spoken language. If we can separate the audio file into segments based on these pauses, we can then process each segment individually and concatenate them to produce the desired output. For this, we made use of python libraries : sounddevice, scipy.io.wavfile, wave, speech-recognition, os, AudioSegment and split on silence.

A technique for extracting word embedding from corpus paragraphs is offered by Doc2Vec. Instead of vector representations of the entire corpus, we can think of these word vectors as paragraph vectors. Doc2Vec is distinct from word2Vec since it uses paragraph id as a distinctive vector. Analyzer will read the pdf file uploaded by HRs and create a dictionary of question and answers, Now create vectors of the question and answer text to compare with the answer given by the candidate

Cosine similarity is a statistic that may be used to assess how similar data objects are, regardless of size. We can measure the similarity between two sentences in Python using Cosine Similarity. Cosine similarity treats each data object in a dataset as a vector. The following is the formula to determine the cosine similarity between two vectors:

$$\text{Cos}(x, y) = \frac{x \cdot y}{\|x\| * \|y\|}$$

- $x \cdot y$ = product (dot) of the vectors 'x' and 'y'.
- $\|x\|$ and $\|y\|$ = length of the two vectors 'x' and 'y'.
- $\|x\| * \|y\|$ = cross product of the two vectors 'x' and 'y'.

On the basis of how similar the answer of the candidate is to the desired answer, the candidate is graded by $(\text{cosine similarity}) \times 100$ percentage.

VI. RESULTS

The need to provide platforms still exists since everyone is looking for opportunities in a fast-paced environment. In a nation like India, where there is so much to contribute, it is crucial to provide everyone with an equal opportunity to demonstrate their knowledge and to give deserving individuals jobs. So, with the booming IT field coming up with solutions, the application of an Interview Analyzer is increasing. In this paper, we built and programmed a working version of Interview Analyzer with a natural language processing approach, that can assist the fair systems hugely in long run. It involves collecting efficient real-time audio data from employees via a microphone and storing it in the form of chunks which are generated after every pause. The given answer is converted to vectors and stored which is used to compare with the expected answers vector to calculate the score. We then calculated cosine similarity between the vectors and the score is given depending on the percentage of similarity calculated.

A. storing Audio Data

Below are the outputs of storing audio by forming chunks of it.

```

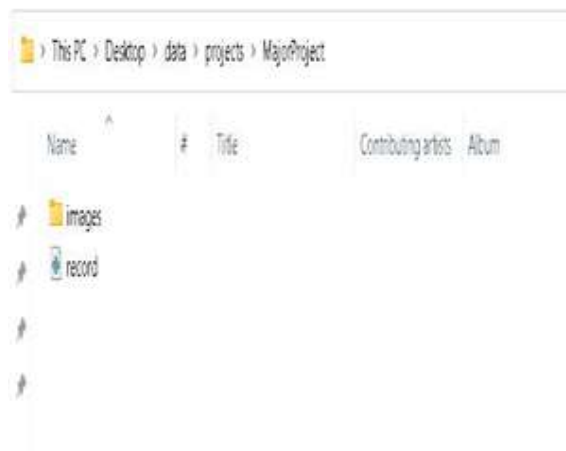
Command Prompt
Microsoft Windows [Version 10.0.22000.978]
(c) Microsoft Corporation. All rights reserved.

C:\Users\DASHRATH ASUTIKAR>cd C:\Users\DASHRATH ASUTIKAR\Desktop\data\projects\MajorProject

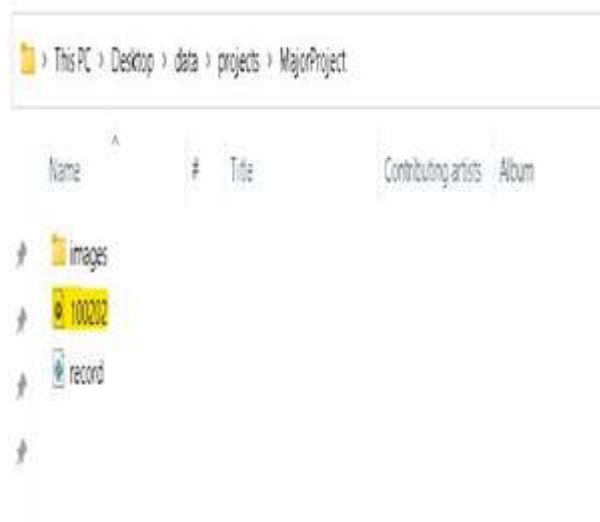
C:\Users\DASHRATH ASUTIKAR\Desktop\data\projects\MajorProject>python record.py
Enter candidate id: 100202
Your recording has been started...
Your recording has been ended.....

C:\Users\DASHRATH ASUTIKAR\Desktop\data\projects\MajorProject>
    
```

(9)



10



11


```
KeyedVectors(vector_size=5, 68 keys)
[-1.176122 -1.0819598 -5.7362795 -2.7640035 2.7603238]
[-0.7079495 -0.7451106 -3.842127 -2.1425748 1.660944 ]
[-0.88649446 -0.7697798 -4.407179 -2.476477 1.9778804 ]
```

C. Calculating score Below is the output of calculating score.

```
Expected answer: A database is an organized collection of data,
Given answer: Database is used to store related data
Your score is : 11.95 %
Your Grade is E2
```

Final Output



pycache	22-11-2022 11:05 PM	File folder
audio-chunks	22-11-2022 11:10 PM	File folder
templates	07-02-2023 08:28 PM	File folder
102_score	07-02-2023 11:29 PM	Microsoft Edge PD...
1001	29-11-2022 09:23 AM	WAV File
1002	29-11-2022 09:27 AM	WAV File
app	07-02-2023 11:28 PM	Python Source File



VII. FUTURE SCOPE

- 1) Analysts progressively rely on video records to analyze the process of understanding. Video uncovers a wide range of comprehensive data sources, that is people use to talk, gesture, gaze, body position, facial expression, movement, and material objects to exchange ideas and information.
- 2) This data is available for repeated viewing and can be an ideal source for capturing and analyzing con-text with that it also provides both breadth (footage that spans weeks or months of activity) and depth (a richly detailed, moment-to-moment interactional record), the traditional qualitative research methods generally do not afford.
- 3) This project can further be enhanced by analyzing video data input for confidence/fluency detection using com-puter vision.

VIII. REFERENCES

- [1] Vatsal, Word2Vec Explained , Jul 29, 2021
- [2] Quoc Le, Tomas Mikolov, Distributed Representations of Sentences and Document, May 2014.
- [3] D Gunawan, The Implementation of Cosine Similarity to Calculate Text Relevance between Two Documents, 2018.
- [4] G`unter R`ohrich, Find Text Similarities with your own Machine Learning Algorithm, Jun 21, 2020.
- [5] Adrien Sieg, Text Similarities: Estimate the degree of similarity between two texts, July 5, 2018
- [6] Masatoshi Nishimura, The Best Document Similarity Algorithm: A Beginner's Guide, Aug 26, 2020.
- [7] Mathew Alexander, Billion-scale semantic similarity search with FAISS+SBERT , Building the prototype for an intelligent search engine, Oct 2020.
- [8] Kay Ellen Ramey, Jaakko Hilpp`o, Elizabeth Dyer, Christina Krist, Qualitative Analysis of Video Data: Standards and Heuristics, June 2016.
- [9] Prerana Das, Kakali Acharjee, Pranab Das, Vijay Prasad, VOICE RECOGNITION SYSTEM: SPEECH-TO-TEXT, November 2015
- [10] Francesco Elia , How to Compute the Similarity Be-tween Two Text Documents?, September 9 2020
- [11] Sentiment Analysis for Interview Transcripts , October 6, 2021
- [12] Hakan Boz and Utku K`ose, Emotion Extraction from Facial Expressions by Using Artificial Intelligence Techniques, Mar 28, 2019.
- [13] <https://www.getsmarter.com/> , Applications of Speech Recognition, Mar, 2019