

Approach Paper

To assess an applicant's suitability for a job based on his video interview

Rajul Shakywar

Indian Institute of Technology, Roorkee, India

Email: rshakywar@ec.iitr.ac.in

Introduction :

In today's competitive world every company is getting a large number of applications for jobs/internships and filtering out most competent candidate by face to face interview for a crucial role is becoming more time consuming and challenging task. The use of AI in recruitment can save time and reduce the HR manager efforts.

Approach to the solution.

Step1.

We have three sets of data from a recorded video interview.

1. Visual Data
2. Audio Data
3. Text Data (Audio → Text)

Step 2.

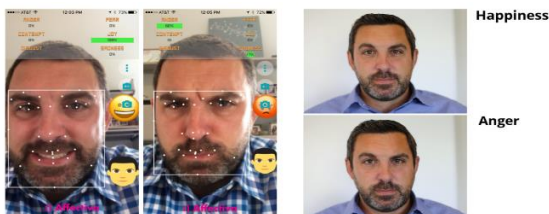
We will process all types of data features extraction.

Visual Data

1. Face Emotions Detection

Methods:

In the first step, we can use **OpenCV** python library for Face detection and in the second step, **CNN** for the classification of emotions. Emotions can be happy, nervous, angry, sad, fear and others.



Some desirable features are

- A smiling face without anger emotions
- Attentiveness based on eye movements.

- No Lips Biting: People sometimes bite their lips when they are worried, anxious, or stressed

2. Gesture and body language analysis

Methods:

EDBL algorithm, Canny algorithm for Edge extraction :

we can extract emotions and state of mind from the human pose, in terms of body gesture, including shoulder and hand. EDBL algorithm can detect emotions even without facial expressions. However, we can also use canny edge extraction technique for body edge detection.



Now our shape is fed into the dataset model to classify the given gesture is desirable or not.

Some desirable features are:

- Sitting in the right posture shows a higher level of confidence
- Fidgeting, tapping of feet
- twirling your hair-conveys a negative impression,
- Well-Dressed up.

Text Data

We can obtain text data from audio data and Speech recognition can be done by using Python API called **SpeechRecognition** or voice assistants like **Alexa**. Moreover, **Google API** is not much accurate for large audio files.

Classification Methods:

We can use **NLP** methods for text data analysis.

ULMFiT

Pre-trained Language Model → Fine Tune on new Dataset → Text classifier

In this method, we fine-tune a pre-trained language model, on a new dataset such a manner that it previously learned is still remembered. We can use it for sentiment analysis and text classification.

Google's Bert

BERT model can be pre-trained with the desirable features and can be fine-tuned with just one additional output layer to perform a wide range of tasks, such as question answering and language inference.

Some desirable Text features

- Use of powerful and desirable keywords such as -
“I am a team-oriented and conscientious person.
“ I am pretty easy going, and I find it easy to get along with the people.”
“ I am very motivated & hardworking and love doing what I do.”
- Avoid Uttering too many fillers, such as 'um', 'like', 'you know' and so on.
- Less Number of negative words like Difficult, Uncomfortable, Impatient etc.

Audio Data

Methods for Feature extraction:

1. **Brain Emotional Learning (BEL)**: It is a Reinforcement learning-based model which shows the function of the limbic system which is responsible for human emotions based on their mammalian brain.
2. **MFCC**: It is a very accurate model for extraction and representation of features of the speech signal.

Some desirable voice sentiments.

- Patient response in a given time limit but should be without long silence which reflects good listening capacity.
- Confident voice
- Tone Analysis (Intonation and Soft Pitch)