

Automated Interview Evaluation using Rule based Chatbot

Dr. S.K Wagh¹, Jitendra Choudhary², Dheeraj Agarwal³, and Laxmikant Hire⁴
^{2,3,4} Student, Department of Computer EngineeringMESCOE, Pune – 01,India
¹Department of Computer Engineering, MESCOE, Pune – 01,India

Abstract: Interview is a way to get employment. We cannot avoid Interviews. Preparing of these interviews are very difficult if it's the first interview and most of the people who have not given any interviews fail to understand the reason for their rejection. We are proposing a system which takes the interview using chat bot and provides a score based on the emotion analysis on the face, sentiment analysis on the text and answer and its evaluation bases on the rules. This score could help people where they are lacking and what need to be done.

Keywords: Chatbot, Interview, Emotion Analysis, Human Computer Interaction, Neural Network.

I. INTRODUCTION

Interview is a process where a conversation occurs in form of the question answers. Interviews are carried out so that the employer can get the best candidates and provide employment. As easy, it sounds, its very difficult to crack an interview without proper preparations and many candidates loose great opportunities due to this. According to Wikipedia, Job In-terview is defined as " Interviews which describe a formal consultation for the purpose of evaluating the qualifications of the interviewee for a specific position. Interviews are seen as a useful tool in assessing qualifications." [1] . The Selection is not only based on your answers but also various other parameters like your way of speaking, your expressiveness. There is a problem in interviews that interviews can be biased based on caste, creed, race, religion, color and so on. This should not be the case as the employer is looking for his requirements and this bias may lead to rejection of perfect candidate. Hence, we are proposing application which can also be a Software as a Service (SAAS) which could be accessed via web and evaluate candidates based on the answers they give, the sentiment of their answers and the emotion during the interview. The conversation agent in our application will be a chat bot which is has interview questions.

Section 1 is Introduction, Section 2 is the related work,Section 3 is about some of the existing systems .Section 4 shows the data sets used, Section 5 is about Proposed System, Section 6 is about utilization of Convolutional neural network for emotion detection in interview process. The Section 7 is the related to Implementation, Section 8 is the Results, Section 9 and 10 are the Conclusion and Future Scope.

II. RELATED WORK

1. Decision Support System for Employee Recruitment Using Multifactor Evaluation Process

The authors put forth a decision support system. The DSS is able to provide Multifactor Evaluation. Results of the research is a decision support system application web-based recruitment of new employees who provide information to the last of the top-ranked candidate new employees as consideration for taking the right decision and is expected to simplify the process of decision the best. [2]

2. Review of integrated applications with AIML based chat bot

The Author reviews the applications of chat bot which are based on AIML(Artificial Intelligence Markup Language). A chat bot can take conversation with humans. AIML is just a xml derivative and is keyword based chat bot. [3]

3. Training Deep Networks for Facial Expression Recognition with Crowd-Sourced Label Distribution

Crowdsourcing has become a widely adopted scheme to collect ground truth labels. [4]. In this paper, the author demonstrate how to learn a deep convolutional neural network (DCNN) from noisy labels, using facial expression recognition as an example. [4] An enhanced FER+ data set with multiple labels for each face image is provided. [4]

4. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications

The paper provides a neural network model which is able to run on the embedded systems with less size and similar accuracy. MobileNet utilizes depthwise separable convolutions which does only a single convolution on each colour channel rather than combining all three and flattening it. [10].

5. Automated Prediction and Analysis of Job Interview Performance: The Role of What You Say and How You Say It

The Author built a model by analyzing 138 recorded inter-view videos of 69 internship-seeking students as they spoke with professional career counselors. [8] Their automated analysis includes facial expressions (e.g., smiles, head gestures), language (e.g., word counts, topic modeling), and prosodic information (e.g., pitch, intonation, pauses) of the interviewees. They derived the ground truth labels by averaging over the ratings of 9 independent judges. [8] Their framework automatically predicts the ratings for interview traits such as excitement, friendliness, and engagement with correlation coefficients of 0.73 or higher, and quantifies the relative importance of prosody, language, and facial expressions. According to their framework, it is recommended to speak more fluently, use less filler words, speak as "we" (vs. "I"), use more unique words, and smile more. [8].

III. EXISTING SYSTEMS

There are many interview evaluation system, as recruitment is very important and difficult for the companies. The following are few of the existing systems:

AutoView by Aspiring Minds (AMCAT)

The following are the advantage and disadvantages of the system:

– Advantages are:

- Evaluates Body language
- Evaluates Emotional Intelligence
- Evaluates Domain Knowledge
- Evaluates Personal Characteristics

– Disadvantages are

- No realtime feedback
- No follow up questions asked only fixed questions

HireVue

The following are the advantage and disadvantages of the system:

– Advantages are:

- Not limited to technical interview, the system can interview on variety of job roles
- Multiple response types (multiple choice, written, coding)
- Scoring is based on 1000+ factors

– Disadvantages are

- No Emotion Detection No realtime feedback
- No follow up questions asked Fixed questions

IV. DATASET

There are two major tasks which the proposed system needs to perform which are as follows:

Emotion detection - Emotion Detection is done by using FER + dataset [4] by Microsoft and over more than 3000 images scraped from the internet. The choice of FER+ dataset was done because it was tagged by 10 different taggers hence we were sure that the image is from the particular category.

Questions Dataset for chat bot - Interview Question dataset was not available, so we created the dataset of our own. The dataset consists of 1928 questions from two broad categories technical and non-technical. These two categories are further divided into sub categories. We created a scrapper using BeautifulSoup library and scrapped the questions and created the dataset. We manually selected the questions for the data sets and assigned categories and subcategories to it. The Fields in the dataset are as follows:

TABLE I
QUESTION DATASET COLUMNS

Dataset columns
Question
Evaluate
Category
Sub_Category
Answer_Keywords

V. PROPOSED SYSTEM

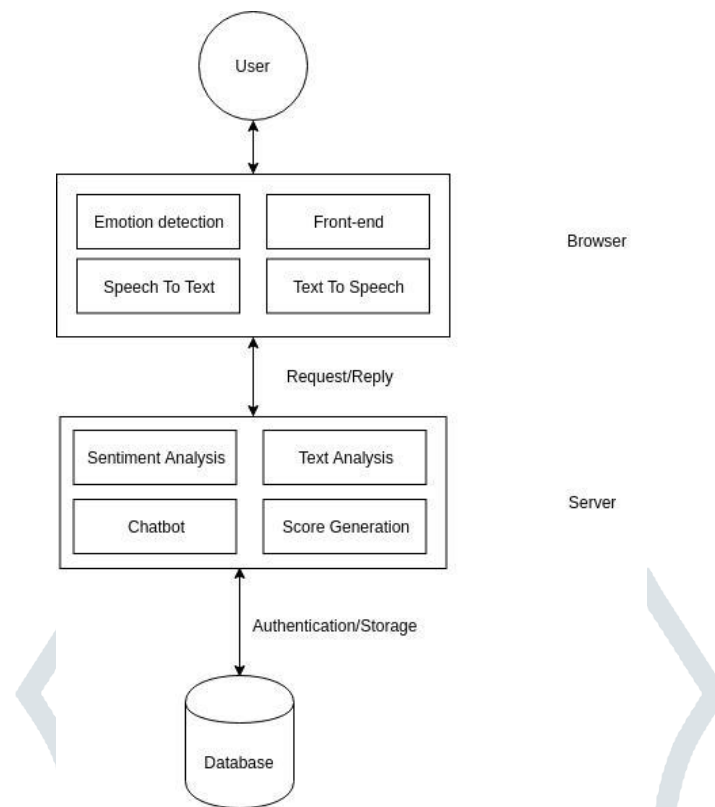


Fig. 1. Proposed System

The proposed system have 3 Tier Architecture - Client, Server, Database.

Client: The Client is simply a Browser. The client is renders front end to the user. There various components which are done on the client side which are as follows:

A) **Emotion Detection** : Emotion detection is done on the client side on the browser using tensorflow.js and videoJS. TensorflowJS allows us to run neural network on the browser itself which kind of reduce the interaction with server. The Emotion detection model is trained with keras and tensorflow in python and then converted to tfjs(json) format which is needed by tensorflowJS.

A) Real time emotion detection is shown to the frontend to the user.

B) **Frontend** : As we are having a client as a browser, the frontend is made using web technologies like HTML, CSS, JavaScript, Ajax and bootstrap. Frontend helps us to represent the interview to the user and also it renders the various visual plots of the analysis in real time.

C) **Text To Speech** : An Interview is a voice based conversation, so we need to inculcate that in our system, so we included JavaScript based web speech API which does speech synthesis in the browser itself. Using web speech synthesis, the user can hear the questions which are asked him or her and answer accordingly.

D) **Speech To Text** : The users also communicates in natural language but our system works on the text hence there is a need to convert speech to text which is also done using web speech API , which have good NLP conversion rate. The transcript is generated from script and then given to chat bot, based on which chat bot ask questions and provides a score.

Server: Server plays a very important role in the application framework as it does the majority of the work is done by the server. The server reply to the request of the client. The component of the Server is as follows:

A) **Sentiment Analysis:** The Sentiment analysis is done on the server side. Each answer given by the user goes through the sentiment analysis. For sentiment analysis we are utilizing nltk based Vader library which does the intensity based sentiment analysis (i.e each word have some kind of intensity associated with it) . The server sends these responses which have the sentiment regarding the answer in positive, negative and neutral. The sentiment score plays a vital role in the score calculations.

B) **Text Analysis:** Text Analysis is done on the answer provided by the user. This involves finding the most repeated words, their frequencies and suggestion based on this. If the speech to text conversation is strong than we can also detect the stop words like um, hmm and so on. The response by this module is the most frequent words with their frequencies, the lexical diversity (i.e ratio of unique words to the total words spoken in the answer).

C) **Chat bot:** Chat bot is a very important component of the system , as it carries out the conversation with the user. Unlike traditional chat bots where the user asks questions and the chat bot answers, the chat bot in this system does the reverse thing(i.e the chat bot ask the question and the user answers it). The chat bot is implemented using AIML (Artificial Intelligence Markup Language) which is an XML derivative. The chat bot have access to over 1600+ questions in various categories like parents,

friends, hobbies, aspiration, food and also on technical questions. This is certainly a rule based chat bot which uses pattern matching to find the intents and ask question on the basis of those intents. [3]

D) **Score Generation** : Each answer is scored based on various parameters like as follows:

- 1) Type of the question (Technical or not technical)
- 2) Emotion Score
- 3) Sentiment Score
- 4) Keywords (If the Type of the question was technical)

The Interview questions are divided into two parts technical and non-technical and the next question is determined by the intent which is determined by the chat bot. If the question is technical then the score calculation is given as follows :

Emotion : 25% Sentiment : 25% Keywords : 25 %

If the question is non-technical then the evaluation is done on the following basis

Emotion : 50% Sentiment : 50%

The steps or scoring are:

Generation of Emotion Score (ES) : The emotion values are scaled using minmax scalar so that they can come in range of 0 to one. Each Emotion has a value associated with which is as follows:

TABLE II
EMOTIONS AND THEIR INTENSITY FOR SCORING

Emotion	Emotion Intensity
HAPPY	1.4
NEUTRAL	1
SUPRISE	- 0.35
SAD	- 0.35
FEAR	- 0.35
DISGUST	- 0.35
ANGER	- 1.4

The formula is given below

$$\left(\sum_{i=0}^{i=8} E_i * ES_i \right) * T$$

Where E_i is Emotion Intensity of Emotion i ,
 ES_i is the Scaled Emotion i ,
 T is the weight age (in technical 0.25 and in nontechnical 0.50)

Generation of Sentiment Score (SS)

Sentiment Score is calculated by using the following the formula

$$\frac{(\sum_{i=0}^{i=3} S_v * W_s) * T}{3}$$

where, S_v is Sentiment value,
 W_v Sentiment Weight.
 T question type (in technical 0.25 and nontechnical 0.50)

Generation of Keyword Score (KS)

$$\frac{NumberofKeywordMatched}{TotalKeywords} * 0.5$$

In Case of non-technical question $KS=0$ Total Score

Score= ES + SS +KS,

where

ES=Emotion Score'

SS=Sentiment Score

KS=Keyword Score

VI. CONVOLUTIONAL NEURAL NETWORK FOR EMOTION RECOGNITION IN INTERVIEW

Convolutional Neural Networks are powerful tool for image classification and recognition. MobileNets are very light-weight compared to VGG-13, VGG-16, ResNET and other such networks because MobileNet utilizes depthwise separable convolutions which does only a single convolution on each colour channel rather than combining all three and flattening it. [10] In MobileNets the depthwise convolution applies a single filter to each input channel. The pointwise convolution then applies a convolution to combine the outputs the depthwise convolution. [10] A standard convolution both filters and combines inputs into a new set of outputs in one step. The depthwise separable convolution splits this into two layers, a separate layer for filtering and a separate layer for combining. This factorization has the effect of drastically reducing computation and model size. [10]

The MobileNet was selected for the simple reason that they can be used on the embedded devices. Running a vgg network on the browser would have had several problems. We applied Transfer learning on the mobilenet with our dataset. We trained the model on 100 epochs with a batch size of 32. The preprocessing steps involves

Get The Frame from the live videostream

Detect the face from the frame (used faceapi.js a open source library for face detection)

Resize the face image to a size of 128*128

Give the image to the MobileNet for prediction

Draw the bounding box with label and send to the video

The operations are done using tensors provided by the TensorFlowJS and hence the above operation can work in realtime

VII. IMPLEMENTATION

The complete application is made using python and web technologies. We understood that the user interface and user experience is very important hence a lot of efforts were taken to make the user experience great. For Backend development we used python based flask framework which is also a micro framework. Flask follows the principle of web development, one drop at a time. The database utilized for authentication and storage was MYSQL. Audio Visualization is done using RecorderJS library. Audio Visualization enables user to visualize their voice, and they can conclude from it how fast or slow and loud and slow they are speaking.

The Data Visualization is done using the ChartJS Library and visualizes the data in the run time. Various types of charts are used to provide the analysis of the interview to the user. The check if the question is asked already is done on the client side. The Emotion detection models runs on the browser itself. This is done to reduce the interaction of the client with the server. The browser supported is only chrome because only chrome browser supports the web speech API as Mozilla Browser to balance the has removed it from their current version.

The front-end is made using CSS and bootstrap with a purple theme.

Following good software development practices we utilized few API instead of reinventing the wheel.

The following are the screen shots of our application

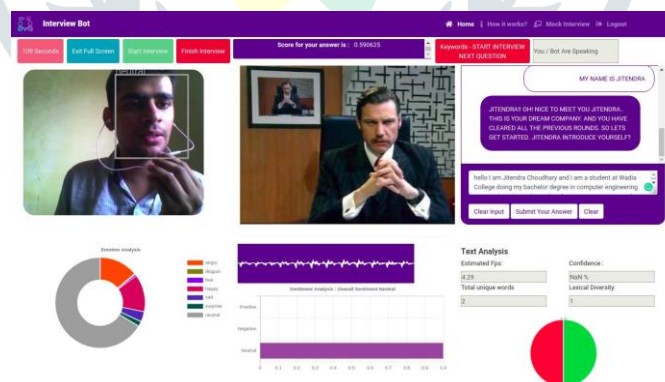


Fig. 6. Mock Interview

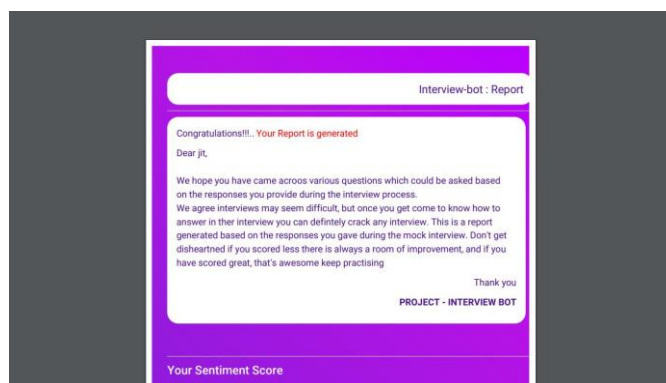


Fig. 7. Report at the end of interview

VIII. RESULTS

As we trained the CNN (Convolutional Neural Network) on the FER+ dataset and 6000 Images scraped from the internet to balance the dataset samples. We used pre-trained Mobile-net ssd model to train for emotion detection. The results which we got are as follows

TABLE III
CONFUSION MATRIX FOR EMOTION DETECTION.

	Neutral	Happy	Surprise	Sad	Anger	Disgust	Fear
Neutral	90.12%	2.06%	2.23%	4.20%	1.02%	0%	0.37%
Happy	3.375%	92.36%	1.7475%	1.22%	1.29%	0%	0%
Surprise	5.97%	3.6%	85.12%	1.62%	1.72%	0%	1.965%
Sad	26.76%	3.445%	1.6075%	61.24%	4.024%	0.9232%	2%
Anger	14.16%	2.93%	3.16%	2.65%	73.3%	3.16%	0.64%
Disgust	9.36%	2.66%	5.14%	2.66%	52.57%	27.49%	0.12%
Fear	3.92%	0%	32.29%	9.27%	5.29%	0%	49.23%

The chatbot uses pattern matching based on the sub category and picks randomly any question which belongs to the subcategory.

The chat bot carries out the conversation as follows:

user :start interview

Bot: OK LET'S START THE INTERVIEW. What is your name ?

user :my name is jitendra

Bot: jitendra!! Oh! Nice to meet you jitendra.. This is your dream company. And you have cleared all the previous rounds. So lets get started. jitendra Introduce yourself?

user :I am jitendra and i am a student doing my graduation in computer engineering as a student i have done 3 internships and many side projects.

Bot: Tell me your internship experience?

user :it was a great company basically a startup This start up was working on product which hugely deals with computer vision and classification and i learned a lot of things.

Bot: Which technology stack did you used in your internship project?

user :I used python for machine learning and android for deploying the model in form of an application. **Bot**: Any projects related to computer vision?

user :I have done projects like human pose estimation and facial keypoints detection and several projects in Computer vision using machine learning.

Bot: What are the different Algorithm techniques in Machine Learning?

user :There are different algorithms like random forest deci-sion tree naive bayes

Bot: How do you define success?

user :i define success when we achieve what we always dreamed of.

Bot: What are you most proud of?

user :I am proud of myself every time i fail at something i always evolve better i think thats very important and my family is always their for me,

Bot: tell me something about your family?

user :my father is business man and my mother is a housewife

my sister is also a house wife she stays in chennai and my

brother is a student

user :end interview.

IX. CONCLUSTION

A chat bot asking you interview questions can surely improve your interview skills and you can come across various questions which could be asked and prepare accordingly. Practice is a key to success in an interview. Emotion detection and sentiment analysis gives a better idea about your speaking style. Hence, we conclude that our proposed system can be used for preparation in interviews.

X. FUTURE SCOPE

This proposed system could be extended and expanded in various ways like stated below:

- 1) The Proposed system can be modelled to a VR based application
- 2) Hand gesture tracking, Extensive Speech Recognition could be added

REFERENCES

- [1] "Interview." [Online]. Available: <https://en.wikipedia.org/wiki/Interview>
- [2] W. Verina, M. Fauzi, F. Nasari, D. H. Tanjung, and J. Iriani, "Decision Support System for Employee Recruitment Using Multifactor Evaluation Process," in 2018 6th International Conference on Cyber and IT Service Management (CITSM), Aug 2018, pp. 1–4. [Online]. Available: 10.1109/CITSM.2018.8674277
- [3] M. S. Satu and M. H. P. and, "Review of integrated applications with AIML based chatbot," in 2015 International Conference on Computer and Information Engineering (ICCIE), Nov 2015, pp. 87–90. [Online]. Available: 10.1109/CCIE.2015.7399324
- [4] E. Barsoum, C. Zhang, C. C. Ferrer, and Z. Zhang, "Training Deep Networks for Facial Expression Recognition with Crowd-Sourced Label Distribution," in ACM International Conference on Multimodal Interac-tion (ICMI), 2016.
- [5] H. Lynda and T. M. E. Amine and B. Farida and B. Tassadit, "E-recruitment support system based on MOOCs," in 2017 International Conference on Mathematics and Information Technology (ICMIT), Dec 2017, pp. 234–238.
- [6] C. Lin, "Behavioral Interview and its Implementation," 2010 3rd In-ternational Conference on Information Management, Innovation Man-agement and Industrial Engineering, Kunming, 2010, pp. 74-76. doi: 10.1109/ICIII.2010.23
- [7] L. Yuanxun and L. Rui, "A Preliminary Analysis on E-interview," 2010 International Conference on E-Business and E-Government, Guangzhou, 2010, pp. 889-891. doi: 10.1109/ICEE.2010.230
- [8] I. Naim and M. I. Tanveer and D. Gildea and M. E. Hoque, "Automated prediction and analysis of job interview performance: The role of what you say and how you say it", 2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), 2015, pp. 1-6. doi=10.1109/FG.2015.7163127
- [9] D. Sokolov and M. Patkin, "Real-Time Emotion Recognition on Mobile Devices", 2018 13th IEEE International Conference on Au-tomatic Face Gesture Recognition (FG 2018), 2018, pp .787-787. doi=10.1109/FG.2018.00124
- [10] Andrew G. Howard and Menglong Zhu and Bo Chen and Dmitry Kalenichenko and Weijun Wang and Tobias Weyand and Marco An-dreetto and Hartwig Adam, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications",CoRR, 2017