2020 Seventh International Conference on Information Technology Trends (ITT). 25-26 November 2020

Interview Data Analysis using Machine Learning Techniques to Predict Personality Traits

Sejal Chopra
Department of Computer Science
Birla Institute of Technology and Science Pilani,Dubai
Dubai International Academic City,Dubai,UAE
sejalchopra97@gmail.com

Siddhaling Urolagin
Department of Computer Science
Anuradha and Prashanth Palakurthi Centre for AI Research
Birla Institute of Technology and Science Pilani,Dubai
Dubai International Academic City,Dubai,UAE
siddhaling@dubai.bits-pilani.ac.in

2020 Seventh International Conference on Information Technology Trends (ITT) | 978-1-7281-8379-4/20/$31.00 ©2020 IEEE | DOI: 10.1109/ITT51279.2020.9320879

***Abstract*— In this paper, we analyze the MIT Interview dataset and find the relation between the various prosodic features (such as intensity, pitch, frequency, etc.) and the likelihood of the person getting a virtuous assessment in the interview. These prosodic features help in rating a person on several grounds such as how engaging or excited or friendly the candidate was. We have demonstrated how selecting only a few of the prosodic features can give better prediction results. This was done by selecting the top features using the 'recursive feature elimination' technique for five personality traits such as 'Engaged', 'Excited', 'Friendly', 'Calm' and 'Speaking Rate'. It was found that for traits such as 'Engaged' and 'Excited', the prosodic features related to intensity play a major role. For personality trait 'Friendly', prosodic features like pitch and duration of pause are more relevant. Similarly, for personality trait 'Calm', prosodic features related to pitch play a major role and so on. Once the top features were selected, we applied three different regression models with tenfold cross validation to determine the best method for predicting these personality traits. These regression models were evaluated by calculating the negative mean squared error, coefficient of determination, etc. Based on the empirical results, decision tree proved to be the best method for predicting the personality traits based on the selected prosodic features.**

Keywords—prosodic features, feature selection, regression, non-verbal behavior, job interviews

I. Introduction

The success of a job interview is dependent on many factors such as the skills of the candidate, his personality traits, how engaging he is and so on. The personality traits of a candidate include not only the facial expressions and verbal content of the speech but also the intonation of the speech. The intonation of the speech plays an equally important role just as the other factors as it can help keep the interviewer engaged and interested in what the candidate has to say. Not only does prosody help us differentiate questions from sentences but it also helps us to recognize a speaker's emotional state, clarify communication and understand sarcasm. These cues within speech help us detect what a person's intention is and without them it can be easy to misinterpret what the person actually means. In face-to-face communication the spoken words form the verbal channel, while everything else represents nonverbal communication. Nonverbal behavior can be perceived aurally (through tone of voice, intonation, and amount of spoken time) and visually (through head gestures, body posture, gaze or facial expressions). Job interviews are an inevitable part of life for all individuals irrespective of the field they are in. It provides an opportunity for the candidate and the potential employer to decide how well your skills align with the company's needs. Job interviews allow the interviewees to get better acquainted with prospective colleagues and obtain information to help them decide if that job is the right one for them. Over the years career counselors and coaches have given out guidelines and advices to succeed in a job interview. Studies in social psychology have shown that smiling, using a louder voice, and maintaining eye contact contribute positively to our interpersonal communications [1], [2]. These guidelines are largely based on intuition, experience, and studies involving manual encoding of nonverbal behaviors on a limited amount of data [9]. During interviews, interviewees must orchestrate their multimodal behaviors, such as speech content, prosody, and nonverbal cues to effectively communicate their qualifications in a limited amount of time [3, 4, 5]. The success or failure of the interviewee's effort is traditionally assessed subjectively by the interviewer, either through a holistic impression or quantitative ratings. The ratings in such situations can be very subjective and can vary from one interviewer to another.

In this paper we have researched on how different prosodic features contribute to a candidate's likeability in a job interview. For this purpose, we have used the MIT Interview dataset. We have also computed which prosodic features contribute the highest to specific personality traits of the candidate. We have further applied several regression models to predict the personality traits that the candidate displays based on his/her prosodic features. This was done by using the top selected prosodic features to predict different personality traits. In the second section of the paper we present the background research that was done in order to dive deep and better understand the topic. The third section of the paper gives an insight into the dataset used for this study. The fourth section of the paper details out the entire methodology followed for the experiments and the results obtained. The last section gives the concluding results of the experiment.

II. Literature Survey

Non-Verbal communication has a major impact on our daily lives. Being adept at picking up nonverbal cues allows the person to relate, engage and establish meaningful conversations with peers, friends, colleagues, etc. It also helps people to build and strengthen their relationships. However, it is important to note that many forms of non-verbal communications are interpreted differently across various cultures. Even a lack of such nonverbal cues can be meaningful and, in itself, a form of nonverbal communication.[6]. In non-verbal communication, subtle aspects of speech like tone, intensity, pitch etc. are categorized as non-verbal paralinguistic communication, and observed in cases where a person is described as 'driving the conversation' or 'setting the tone' of the conversation [7].

In a job interview, the applicant's nonverbal behavior has a remarkable impact on the hiring decision. For instance, Imada and Hakel showed that applicants who use more immediacy nonverbal behavior (i.e., eye contact, smiling, body orientation toward interviewer, less personal distance) are perceived as being more hirable, more competent, more motivated, and more successful than applicants who do not [4].

Nonlinguistic vocal signaling can tell a lot about a person. For example, some people are confident when they speak and can drive the conversation. Such people are able to dominate the tone and direction of the conversation and are considered highly influential. Such skills are highly in demand for sales persons and social connectors [9]. In 1972, Edward Sapir symbolized non-verbal behavior as "an elaborate and secret code that is written nowhere, known by none, but understood by all" [10]. Hall et al. [11] studied the nonverbal cues in doctor-patient interaction and showed that doctors who are more sensitive to nonverbal skills received higher ratings of service during patient visits. Ambady et al. [12] studied the interactions of teachers with students in a classroom and proposed a framework for predicting teachers' evaluations based on short clips of interactions. The draw back with these frameworks is that they were reliant on manually annotating behavioral patterns. Manually doing this is time consuming, arduous, prone to errors and not an efficient approach when dealing with a huge amount of data.

III. Interview Data Analysis

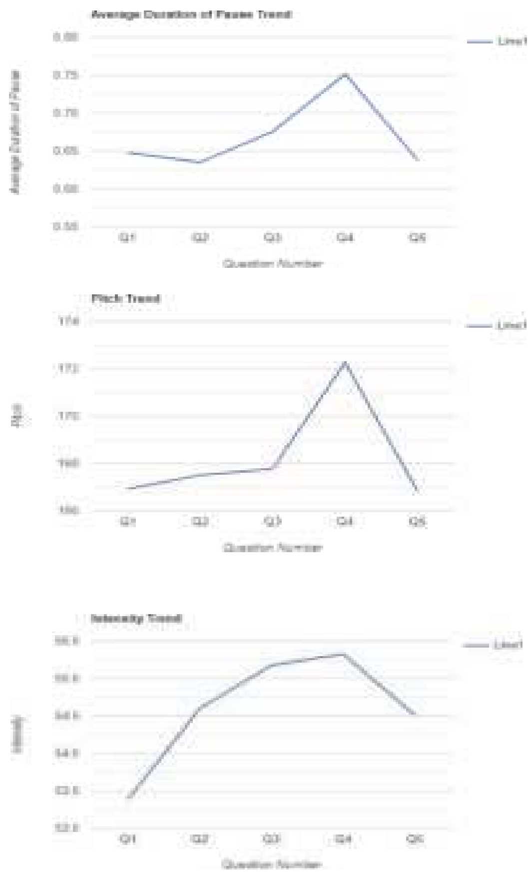
For the purpose of this research we used the MIT Interview Dataset [13], which comprised of 138 audio-visual recordings of several interviews with students who were seeking internships at MIT. There are 138 audio files of the interview with a total duration of 10.5 hours. The average time per audio is 4.7 minutes per interview. Initially 90 MIT students had decided to take part in the mock interviews conducted. The interviews were conducted by two MIT career counselors, each of whom had over five years of experience. For each participant two rounds of interviews were conducted. After the data collection, 69 out of the 90 initial participants permitted the use of their interview recordings for the research. Out of these participants,26 were males and 43 were females. Table 1 details out the five questions that were asked to all candidates.

Apart from the career counsellors, Amazon mechanical turk workers were also used to rate the interview performance to overcome any kind of biases. Apart from being less affected by bias, the mechanical turk workers could pause and replay the video, thus allowing them to rate more thoroughly. The Turkers' ratings are more likely to be similar to "audience" ratings, as opposed to "expert ratings". Looking closely at the data, we could make some observations that helped us to understand the dataset better. The data consists of information on 58 prosodic features for each participant for each of the five questions. The important prosodic features include pitch information, vocal intensities, characteristics of the first three formants, and spectral energy. These features are considered important while talking about one's social behavior [14].

TABLE 1. LIST OF QUESTIONS ASKED DURING INTERVIEW

Question Number	Question
1	So please tell me about yourself.
2	Tell me about a time when you demonstrated leadership.
3	Tell me about a time when you were working with a team and faced a challenge. How did you overcome the problem?
4	What is one of your weaknesses and how do you plan to overcome it?
5	Now, why do you think we should hire you?

Fig. 1 displays the overall pattern for few of the prosodic features for each of the questions that can be observed across all participants. It appears from these figures that the participants took the maximum pauses while talking about their weaknesses while their spectral energy seems to be the highest while talking about their leadership skills. The pitch and the intensity of the speech also is highest during the fourth question. The energy and speak rates for all candidates is comparatively low while talking about the challenges faced by them.



978-1-7281-8379-4/20/$31.00 ©2020 IEEE

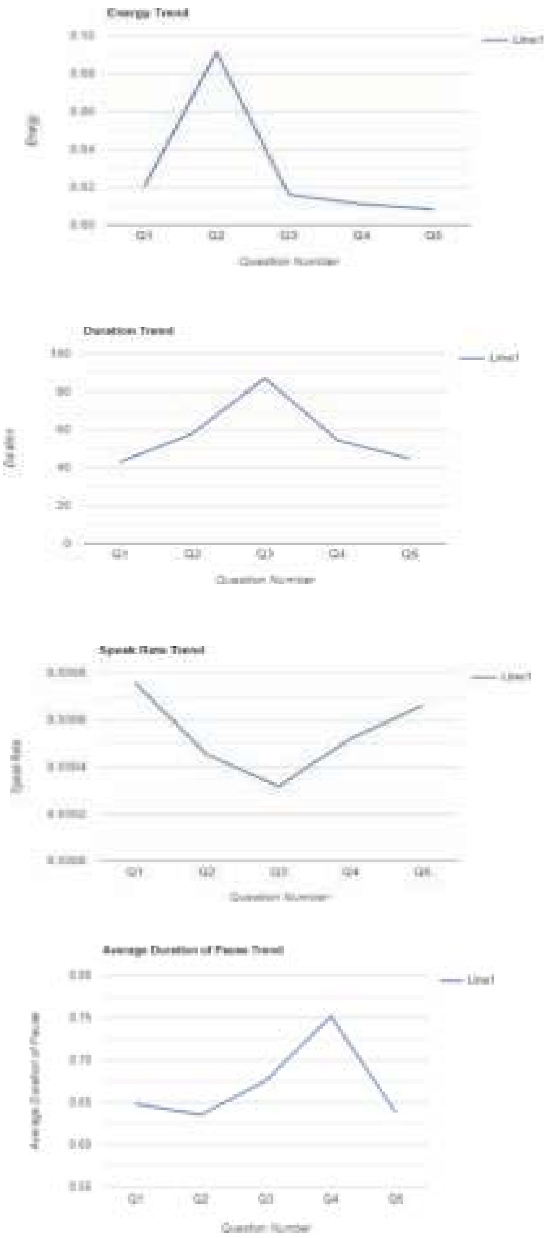


Fig. 1. Trend of prosodic features across different questions

IV. Experimental results

This section deals with predicting the rating of the interviewee based on his/her prosodic features. In order to get the best possible results, we first performed feature selection for traits such as 'Engaged', 'Excited', 'Speaking Rate', 'Friendly' and 'Calm'. Irrelevant or partially relevant features can negatively impact model performance. The choice of feature selection depends on the input and output variables of the model we aim to build. When it comes to the implementation of feature selection, numerical and categorical features are to be treated differently. As in our case, we have numerical input and output data, we have chosen the 'Recursive Feature Selection (RFE)' method to select only relevant features. RFE is a wrapper-type feature selection algorithm. This means that a different machine learning algorithm is given and used in the core of the method, is wrapped by RFE, and used to help select features. RFE works by searching for a subset of features by starting with all features in the training dataset and successfully removing features until the desired number remains. We have selected 'Decision Tree Regressor' as the underlying method for our RFE. We have used 'negative mean squared error' to compare the performance at each step as the number of features are varied from one to fifty-eight. From the following figure (Fig.2), it can be seen that the error rate is lesser on taking a subset of the features to taking all features into consideration. It clearly indicates that the top selected prosodic features based on least error are a good indicator to assess the personality trait.

Fig.2 shows the variation of negative mean squared error when the number of features are varied from one to fifty eight for different personality traits. As per the figure, it is clear that for the personality trait 'Friendly', it is optimal to select only six features instead of selecting all for better prediction. Similarly for the trait 'Calm', it is best to select thirteen features to get the best results and so on. Reducing the number of features will help improve the accuracy of prediction and help reduce the computational cost of modelling. This is because features selection primarily focuses on removing non-informative and redundant features. The details for each prosodic feature used for the model is presented in the paper [14].

For traits such as 'Engaged' and 'Excited', the prosodic features related to intensity(such as intensityQuant, diffIntMaxMode, etc) and pitch (such as mean_pitch, pitch_abs, etc) play a major role. For personality trait 'Friendly', prosodic features like pitch and duration of pause are more relevant. Similarly, for personality trait 'Calm', prosodic features related to pitch play a major role and so on. These results are also rather intuitive and rational. Table 2 shows the list of the top prosodic features selected.

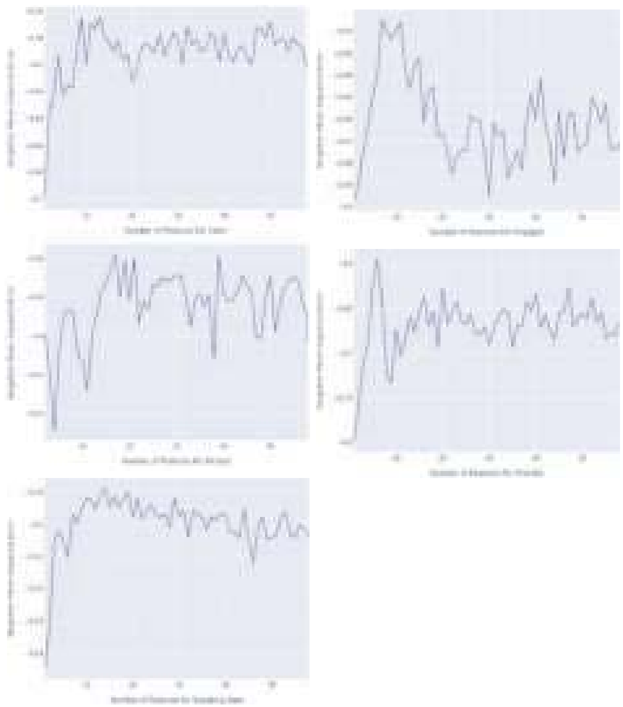


Fig. 2. Variation of mean squared error on change in number of features

TABLE 2. TOP PROSODIC FEATURES SELECTED FOR EACH PERSONLAITY TRAIT

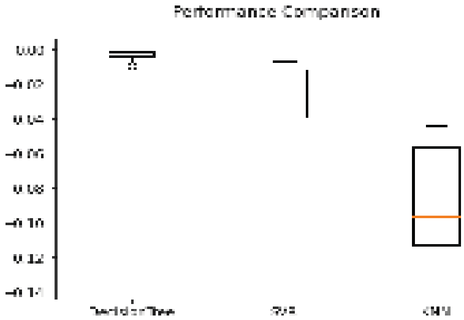
Friendly	Engaged	Excited	Speaking Rate	Calm
mean_pitch	energy	Energy	power	diffPitchMaxMean
avgBand1	intensityQuant	mean_pitch	pitch_abs	energy
f2meanf1	fmean3	pitch_sd	pitchUvsVRatio	Time:8
f1STD	f3meanf1	diffPitchMaxMin	iDifference	diffPitchMaxMin
maxDurPause	f2STDf1	intensityQuant	intensityQuant	diffIntMaxMode
numFall	jitterRap	diffIntMaxMode	diffIntMaxMode	fmean3
	diffIntMaxMin	avgVal3	avgBand2	f2meanf1
		avgBand1	avgBand3	f3STD
		fmean2	f2meanf1	f2STDf1
		f2meanf1	speakRate	percentUnvoiced
		jitterRap	numPause	PercentBreaks
		speakRate	maxDurPause	TotDurPause:3
		numPause	numRising	MaxRising:3
		avgDurPause	numFall	
		AvgTotFall:3		
		numRising		
		numFall		

Using these selected features, we have further quantified the role of each selected prosodic feature for the interview personality trait prediction. To achieve this, we have implemented three regression models and analyzed the results to select the best possible method to achieve accurate prediction results. The three selected regression models used were 'Decision Tree', 'K Nearest Neighbor (KNN)' and 'Support Vector Regression (SVR)' with tenfold cross validation. The results from each of these methods were evaluated using negative mean squared error (NMSE), coefficient of determination(R2) and negative mean absolute error (NMAE). Table 3 below shows the results from each of the regression models.

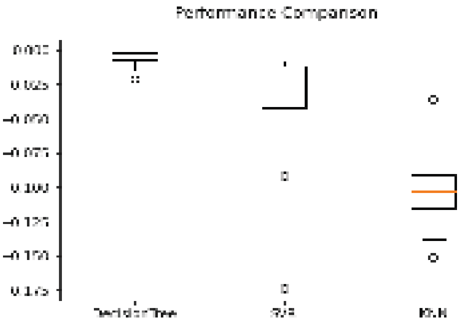
TABLE 3. EVALUATING THE THREE REGRESSION MODELS

Personality Trait	Score	Decision Tree	SVR	KNN
Friendly	NMSE	-0.003186	-0.028199	-0.08805
	R2	0.990602	0.920082	0.761654
	NMAE	-0.02172	-0.07632	-0.18493
Engaged	NMSE	-0.006668	-0.043788	-0.10292
	R2	0.982181	0.875338	0.660416
	NMAE	-0.022472	-0.093334	-0.20381
Excited	NMSE	-0.008871	-0.13286	-0.25644
	R2	0.989537	0.735915	0.43334
	NMAE	-0.038341	-0.161567	-0.32779
Speaking Rate	NMSE	-0.006531	-0.045113	-0.06977
	R2	0.941746	0.707695	0.488955
	NMAE	-0.011643	-0.072134	-0.14799
Calm	NMSE	-0.010603	-0.056309	-0.13305
	R2	0.976252	0.825554	0.585808
	NMAE	-0.023554	-0.113506	-0.24267

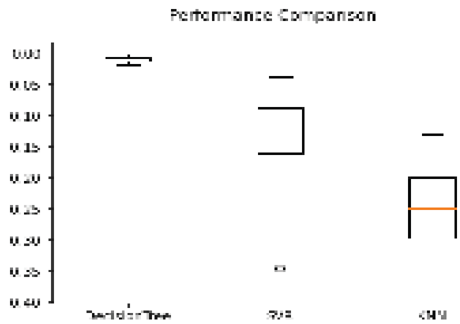
From Table 3, it can be seen that 'Decision Tree' is the best choice from all the three regression models for prediction. The same can be visualized by plotting the negative mean squared error on a box plot for each of the personality trait as can be seen in Fig.3.



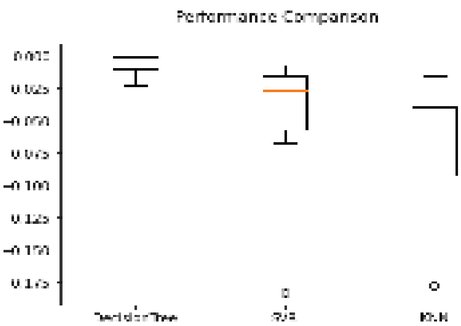
(a)



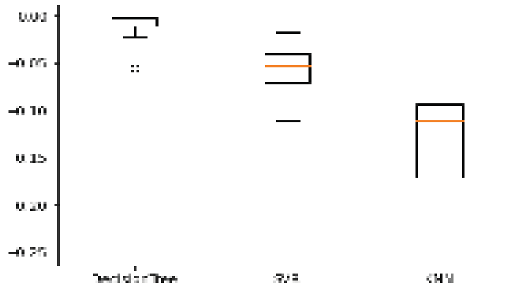
(b)



(c)



(d)



(e)

Fig. 3. Box plot to evaluate the regression models using negative mean squared error (a) Friendly (b) Engaged (c) Excited (d) Speaking Rate (e)Calm

V. Conclusion

Best organizations succeed not because of the talent they have, but because they have the Right People. Behavioral interviewing is a more accurate, cost-effective and flexible method of selection to help companies recruit talents from a pool of applicants and maximize their benefits [14]. In this paper, we present an automated prediction framework for quantifying social skills for job interviews. The proposed model shows reassuring results and predicts human interview ratings. Several traits such as 'Engagement', 'Excited' and 'Calm' were predicted using several prosodic features. Fig. 4 shows the diagrammatic summary of the procedure followed for this experiment. By observing the empirical analysis and results in the earlier section, it is clear that selecting top prosodic features based on recursive feature elimination gives better performance rather than selecting all fifty-eight features. Hence it an enhancement to the work [14]. Once the top prosodic features have been selected for each of the personality trait we have applied different regression models to find the best way to predict these personality traits. Table 3 and Fig. 3 also show that decision tree is the most appropriate choice to predict different personality traits in this case as it gives the least error in all scenarios. Hence it is the best method out of the three to quantify the different personality traits. After decision tree, support vector machine gives the next best performance which is then followed by k nearest neighbor regression model in terms ofperformance. It should be noted that personality is shaped by both genetic and environmental factors. The culture in which you live is one of the most important environmental factors that shapes your personality. Culture is transmitted to people through language as well as through the modeling of culturally acceptable and nonacceptable behaviors that are either rewarded or punished [16]. It looks as if that there are both universal and culture-specific aspects that account for variation in people's personalities. Therefore, for certain individuals some behavioral interpretations might represent a different meaning. This opens gates for future research possibilities so that different cultural contexts can also be accommodated for. Personality traits recognition in interviews has the potential to change the entire landscape of recruitment by adding a layer of artificial intelligence on top of the social interactions that are in place in the current system. However, one has to take caution and ponder on the cultural implications of such an approach to recruitment. The question of consent and privacy comes to mind whether the candidates that come to the interview have been asked for consent or not. Even if consent is given, knowing that an algorithm is analyzing your emotions through your facial expressions is bound to make one self-conscious. Furthermore, it is certain that the power to make this approach the norm is in the hands of the recruiters and the candidates. Such impacts to our culture and society need to be discussed for the technology to mature and to be seen as a positive tool in recruitment.

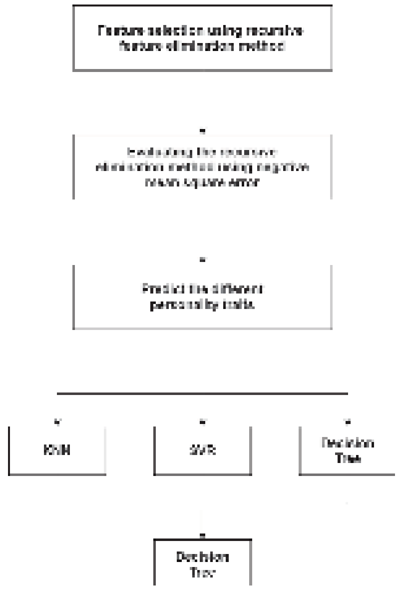


Fig. 4. Procedure followed for the experiment

References

[1] A. I. Huffcutt, J. M. Conway, P. L. Roth, and N. J. Stone,"Identification and meta-analytic assessment of psychological constructs measured in employment interviews." J. Appl. Psychology, vol. 86, no. 5, 2001, Art. no. 897.

[2] T. Macan, "The employment interview: A review of current studies and directions for future research," Hum. Resource Manag. Rev., vol. 19, no. 3, pp. 203–218, 2009.

[3] R. J. Forbes and P. R. Jackson. Non-verbal behaviour and the outcome of selection interviews. Journal of Occupational Psychology, 53(1):65–72, 1980.

[4] A. S. Imada and M. D. Hakel. Influence of nonverbal communication and rater proximity on impressions and decisions in simulated employment interviews. Journal of Applied Psychology, 62(3):295–300, 1977.

[5] C. K. Parsons and R. C. Liden. Interviewer perceptions of applicant qualifications: A multivariate field study of demographic characteristics and nonverbal cues. Journal of Applied Psychology, 69(4):557–568, 1984

[6] Nonverbal communication. (2015, April 6). GoodTherapy.org Therapy Blog. https://www.goodtherapy.org/blog/psychpedia/nonverbal-communication

[7] A. Pentland, "Social Dynamics: Signals and Behavior", ICDL, San Diego, CA, Oct 20-23,2004

[8] Gladwell, M.: The Tipping Point: How Little Things Can Make a Big Difference. Little Brown, New York (2006)

[9] E. Sapir, Selected Writings of Edward Sapir in Language, Culture and Personality, vol. 342. Berkeley,CA, USA:Univ. California Press, 1985.

[10] J. A. Hall, D. L. Roter, D. C. Blanch, and R. M. Frankel, "Nonverbal sensitivity in medical students: Implications for clinical interactions," J. General InternalMed., vol. 24, no. 11, pp. 1217–1222, 2009.

[11] N. Ambady and R. Rosenthal, "Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness," J. Personality Social Psychology, vol. 64, no. 3, 1993, Art. no. 431.

[12] M. E. Hoque, M. Courgeon, J.-C. Martin, B. Mutlu, and R. W. Picard,"MACH: My automated conversation coach," in Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput., 2013, pp. 697–706.

[13] R. W. Frick, "Communicating emotion: The role of prosodic features." Psychological Bulletin, vol. 97, no. 3, 1985, Art. no. 412.

[14] C. Lin, "Behavioral Interview and its Implementation," 2010 3rd International Conference on Information Management, Innovation Management and Industrial Engineering, Nov. 2010.

[15] I. Naim, M. I. Tanveer, D. Gildea and M. E. Hoque, "Automated Analysis and Prediction of Job Interview Performance," in IEEE

Transactions on Affective Computing, vol. 9, no. 2, pp. 191-204, 1 April-June 2018, doi: 10.1109/TAFFC.2016.2614299.

[16] H. C. Triandis and E. M. Suh, "Cultural Influences on Personality," Annual Review of Psychology, vol. Vol. 53:133-160 , Feb. 2002.