Sinhgad Institutes

The 7th Post Graduate Conference of Computer Engineering (cPGCON-2018)

Intelligent Chatbot For Conducting Automated HR Interviews

Paper ID: 198

Track: Data Mining and Information Retrieval

Presented by: Mr. MILIND THOMBRE

Guided By: Prof. Dr. SIDDHIVINAYAK KULKARNI

College Name: MITCOE

College Code: 38

Contents

Introduction

- Usually the last leg of any selection process is the Human Resources (HR) interview. Sometimes these are also mandated when someone is put up for promotion to gauge suitability for the position the candidate would assume.
- An increasingly large number of these generic interviews take place across industry and academia and basically between any candidate and potential employer, including Government.
- We observe issues with the quality and associated costs, both direct and indirect, of such interviews caused by human nature and the inherent variability in the skill and experience of the interviewer. To name a few we observe that several times such
- Interviews are not standardized, of poor quality and the content is variable and not addressed fully

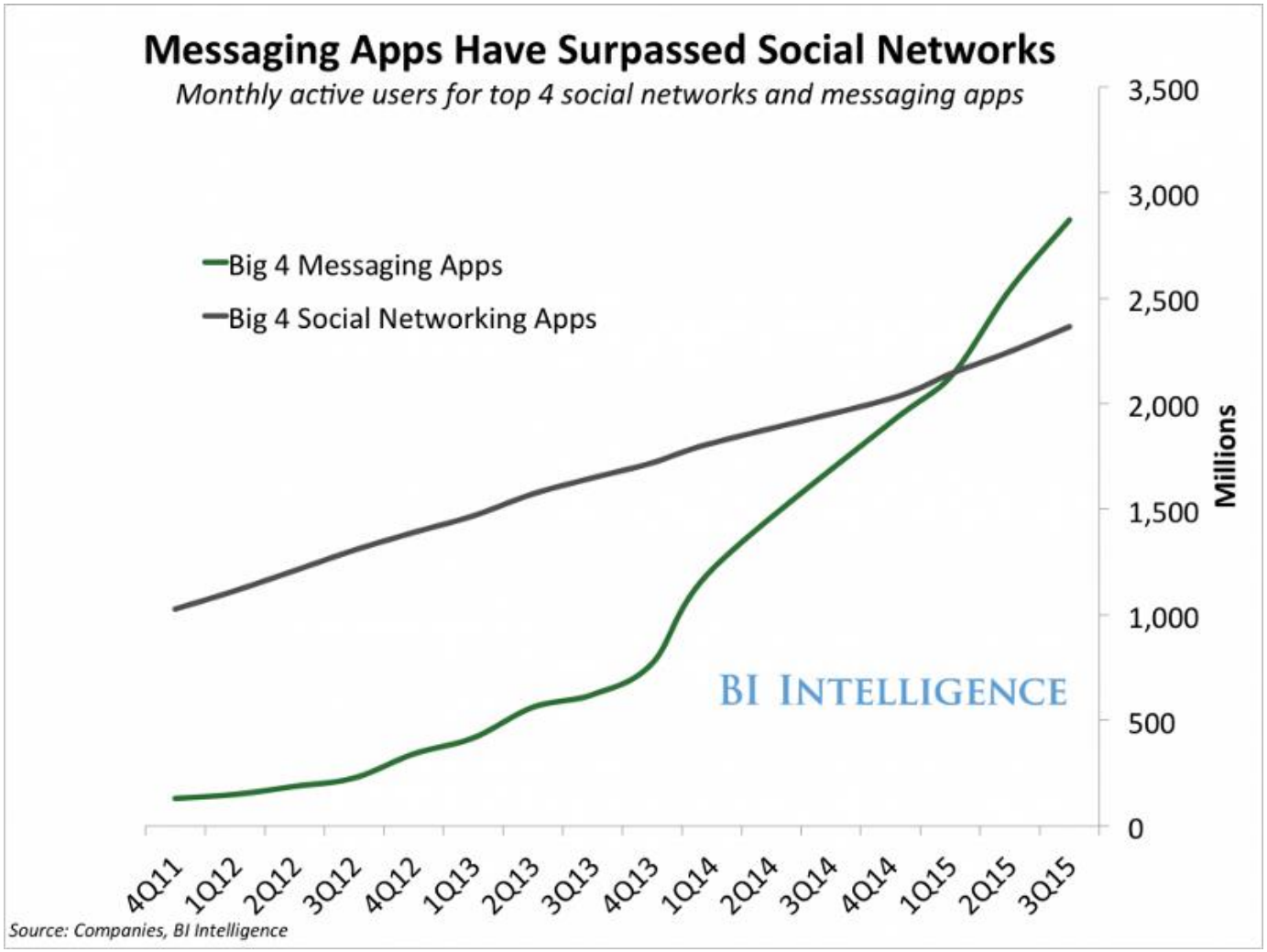
Introduction..

- Presence of bias of the interviewer, non-scalability of the process due to bottlenecks caused by skilled human interviewer availability.
- We also note that the recruiting organizations may be exposed to liability due this defective and human oriented subjective process.
- We note that Machine learning once applied successfully will constantly improve the process of the interview as more training data becomes available and feedback loops of current performance and outcomes versus past interview performance are closed by our Intelligent Chatbot System

Motivation

- Issues observed occurring due to human subjectivity are persistent in nature (bias, subjectivity, opinion etc.) and have not yet been addressed via technology yet even though they were first studied in the year 1911.
- We are motivated by the need for greater justice in the interview process.
- It is generic in nature and once automated, it will be applicable across the spectrum of industry, government as well as academia. This is unlike the technical interview which would have a much narrower market size.
- It is far more subjective and challenging to automate and hence an intellectually stimulating topic.

Motivation


Messaging Apps Have Surpassed Social Networks
Monthly active users for top 4 social networks and messaging apps
Big 4 Messaging Apps
Big 4 Social Networking Apps
BI INTELLIGENCE
3,500
3,000
2,500
2,000
1,500
1,000
500
0
Millions
4Q11
1Q12
2Q12
3Q12
4Q12
1Q13
2Q13
3Q13
4Q13
1Q14
2Q14
3Q14
4Q14
1Q15
2Q15
3Q15
Source: Companies, BI Intelligence

Our Contribution

❖ A Novel Idea and Solution!

Current Issues:

- Personal Bias of the interviewer.
- Accuracy and standardization of the recruitment process.
- Time spent on the process by HRA as well as associated cost

Solution:

The chatbot software addresses a specific segment of recruitment process to help HR Administrators in collecting and analyzing various personal and professional information from the job seekers/promotees in order to estimate their

- fitment in the company and provide a
- hiring decision as well as inputs for considering the
- compensation and benefits to be offered to the candidate.

❖ 2. Algorithm/Model

Literature Survey

- Thombre et al. [1] 2018 note that bias is omnipresent in interviews and that the issue persists historically and to this day.
- In a 1982 Study [2] Arvey et al. note the low reliability in the past of personnel interviews and a weak correlation with job performance and lack of standardization as the same candidates presented to different hiring managers produced dramatically different results!
- Wagner [3] 1949 stressed that free-form or open-ended techniques and recommended that the interview may be useful in only 3 situations:

  1.When rough screening in required. (Extremely large number of initial applicants)

  2 When number of applicants is too small to warrant structured testing and other methods.

  3 When certain traits required for the job at hand may be readily accessed by the interviewer.

An Illustration

- The interview Questions and recording of the candidate Answers via the Chatbot.
- The conversion of the text transcript of the interview to data, its analysis against a standard set of responses.
- Application of Text mining and Machine Learning Techniques to the interview transcript to convert the textual data to meaningful, quantifiable, and actionable information.
- Presentation of the analytics report to the hiring manager to get a final decision from him/her.

Our Proposed Approach/Mathematical Model

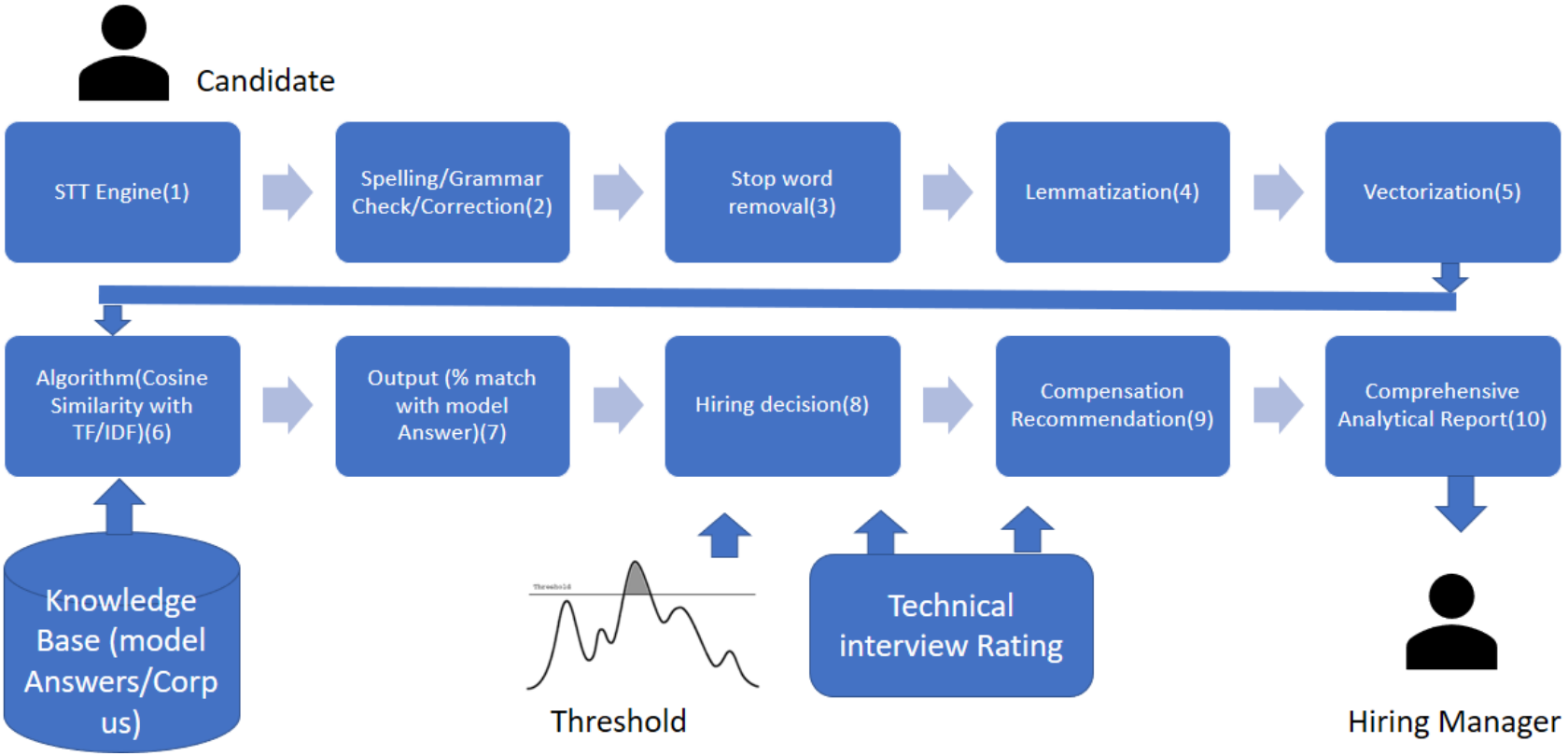
- The central outline of the proposed algorithm is as follows.
  - Step 1: We have a question engine which puts out a question from sorted list of questions
  - Step 2: Record the answer and Convert from Speech to text
  - Step 3: Spelling and Grammar correction of the recorded text
  - Step 4: Remove stop words
  - Step 5: Lemmatization
  - Step 6: Vectorization

Our Proposed Approach/Mathematical Model

- Step 7: Compute Cosine Similarity with the corresponding model answer from Knowledge Base
- Step 8: Repeat steps 1-7 till All questions are answered.
- Step 9: Output % overlap with Model Answers
- Step 10: Repeat steps
- Step 11: Recommend Hiring Decision (Y/N)
- Step 12: Compensation Recommmendation after taking into account Technical Interview Rating
- Step 13: Comprehensive Analytical Report to Hiring Manager

Proposed Approach

Detailed System Architecture



Methodology of Evaluation

- We plan to Implement this system in parallel with the Existing Paper and Pencil System of evaluations in an organization.
- We plan to compare our System Decisions with those arrived at by human Hiring Managers.

Metrics of Evaluation

Z-Score:

$$z_i = \frac{x_i - \bar{x}}{s}$$

Data Set

- Source 1: Answers obtained from Experienced Hr professionals (20+ years in industry)

Attributes: Question, Model Answer, Actual Answer

- Source 2:A crowdsourced website for HR interview Answers www.indiabix.com

Attributes: Question, Answers (multiple), Crowdsourced Rating of answers

Our Observations

- Bias In human Interview Process is omnipresent!
- Inconsistent coverage of the material can be eliminated by using automated software/machine.
- Reliability and Standardization Improve with Automation

Conclusion

- Reliability of the interview process can be increased as a single interviewer (machine) conducts ALL HR interviews.
- Standardization is easily achieved using the same computer model for various candidates.
- In our opinion, Manipulative behaviour such as impression management can be eliminated using bias-less and emotionless machines
- The automated software can be made blind to factors such as race, gender, age, and physical disabilities of the candidate, amongst other things and focus on competence alone

Future Work

- Use newer Techniques Such as Deep Learning (using TensorFlow) once more Labelled Data becomes Available.
- Compare and contrast the results of our approach with Deep Learning Approach

References

[1] Milind Thombre, Siddhivinayak Kulkarni, Survey Paper, The case for an automated HR evaluation tool, 2018. (yet to be published)
[2] Journal of Industrial Psychology, Arvey et al. 1982
[3] Journal of Industrial Psychology, Wagner et al. 1949
[4] Journal of Industrial Psychology, Mayfield et al. 1964
- [5] Journal of Industrial Psychology, Schmitt et al. 1976

[6] Journal of Industrial Psychology, Landy et al. 1976
[7] Journal of Industrial Psychology, Kalin and Rayco 1978.
[8] Journal of Industrial Psychology, Imada an Hakell 1977
[9] Journal of Industrial Psychology, Marchese 1993.
[10] Journal of Industrial Psychology, Stevens and Kristoff 1995.
[11] Journal of Industrial Psychology, Tullar 1989.
[12] Journal of Industrial Psychology, Pingitore et al. 1994
[13] Journal of Industrial Psychology, Harris, 1989.
[14] Bamman, D., O'Connor, B., Smith, N.A.: Learning latent personas of lm characters. In: Proceedings of the Annual Meeting of the Association for Computational Linguistics (ACL). p. 352 (2014)

References

[15] Bird, S.: Nltk: the natural language toolkit. In: Proceedings of the COLING/ACL on Interactive presentation sessions. pp. 6972. Association for Computational Linguistics (2006)
[16] Bruijnes, M., op den Akker, R., Hartholt, A., Heylen, D.: Virtual suspect william. In: International Conference on Intelligent Virtual Agents. pp. 6776. Springer (2015)
[17] Duplessis, D., Letard, V., Ligozat, A.L., Rosset, S.: Joker chatterbot rewochat 2016-17 shared task chatbot description report. In: RE-WOCHAT: Workshop on Collecting and Generating Resources for Chatbots and Conversational Agents-Development and Evaluation Workshop Programme (May 28th, 2016) p. 45
[18] Fialho, P., Coheur, L., Curto, S., Claudio, P., Costa, A., Abad, A., Meinedo, H., Trancoso,I.: Meet edgar, a tutoring agent at monserrate. In: ACL (Conference System Demonstrations). pp. 6166. Citeseer (2013)
[19] Filippova, K.: Multi-sentence compression- Ending shortest paths in word graphs. In: Proceedings of the 23rd International Conference on Computational Linguistics.pp. 322330. Association for Computational Linguistics (2010)
[20] Kubon, D., Hladk, B.: Politician re-wochat 2016 - shared task chatbot description report. In: RE-WOCHAT: Workshop on Collecting and Generating Resources for Chatbots and Conversational Agents-Development and Evaluation Workshop Programme (May 28th, 2016). p. 43

References

[21] Le, Q., Mikolov, T.: Distributed representations of sentences and documents. In: Proceedings of The 31st International Conference on Machine Learning. pp. 1188 1196 (2014)
[22] Li, J., Galley, M., Brockett, C., Gao, J., Dolan, B.: A persona-based neural conversation model. arXiv preprint arXiv:1603.06155 (2016)
[23] McCrae, R.R., Costa Jr, P.T.: A ve-factor theory of personality. Handbook of personality: Theory and research 2, 139153 (1999)
[24] Owoputi, O., O'Connor, B., Dyer, C., Gimpel, K., Schneider, N., Smith, N.A.: Improved part-of-speech tagging for online conversational text with word clusters. Association for Computational Linguistics (2013).
[25] Rehurek, R., Sojka, P.: Software Framework for Topic Modelling with Large Cor-pora. In: Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks. pp. 4550. ELRA, Valletta, Malta (May 2010), http://is.muni.cz/publication/884893/en.
[26] Rushforth, M., Gandhe, S., Artstein, R., Roque, A., Ali, S., Whitman, N., Traum, D.: Varying personality in spoken dialogue with a virtual human. In: International Workshop on Intelligent Virtual Agents. pp. 541542. Springer (2009)

Thank You