Artificial Intelligence Candidate Recruitment System using Software as a Service (SaaS) Architecture

Divyanshi Mishra[1], Sushant Shekhar[2]

[1]*Student, Department of Computer Science and Engineering, SRM Institute of Science and Technology, Tamil Nadu, India*

[2]*Student, Department of Computer Science and Engineering, SRM Institute of Science and Technology, Tamil Nadu, India*

---***---

Abstract — The motive of this research is to utilize the machine intelligence concepts and the algorithms to help firms recruit potential employees. Recruiting employees having core competencies that are maximally aligned with those of the company is the need of every company to come out with flying colors in future. The Chatbot can test the employees in both technical as well as non-technical fields. During the recruitment process, in order to judge the candidate's attitude taking a particular string i.e. positive, negative, or neutral, the disclosed method makes use of sentiment analysis. For measuring the similarity between two strings a Dice coefficient is used. The benefit of this Chatbot is to save time and money of the firm which they spend in hiring the candidates for their organization. To gladden and aid the employees, there is a feature that they can attend the interview from any place at any time according to their convenience. We programmed the Chatbot using node.js, python. It even solves user's grievances, which leads to customer satisfaction to a greater extent. The architecture is highly ductile, robust, vigorous and extensible. The system is implemented on AWS public cloud.

Key Words: Knowledge based systems, Web design, Data warehouses, Web services, Text mining

1. Introduction

Every company's hiring process may vary to a large extent, but there may be one common theme which is followed by every company irrespective of the size of an organization and that is the interview process that carries a majority of the weight in the hiring process of a candidate. As a standard practice, companies conduct various rounds for each vacancy. But practically, it may not work. The job market as we all know is tough, and despite the best efforts put in by the recruiters for creating a stellar hiring process, any compromise in the quality of candidates hired during the interview may dilute the overall quality of the knowledge capital within the company. Further a bad interview process might hurt the business ethics of a company to a great extent. So instead if we eradicate the uncertainties from the interview process and through an online interviewing process accurately gauge the skills of the candidates in a particular or specific domain, the risk posed by traditional methods of recruitment process may get mitigated to a large extent and the resources saved therein can be channelized for other endeavors for the benefit the organization for a longer term. We did a survey with over 300 HRs all around the globe.

The goal of the survey was to figure out what traits do the HRs look after in a candidate to find the perfect for a particular job.

The Survey had several articles like Body Language, Endurance, Trustworthiness, Ambition, Culture Fit, Previous Industry Experience, Co-Working abilities, Behavioral Questions, Patience, Skill Level and Psychometric Score.

We found out that most important factors were Culture Fit, Previous Industry Experience, Behavioral Questions, Skill Level and Psychometric Score. The Pie Chart below illustrates the same.
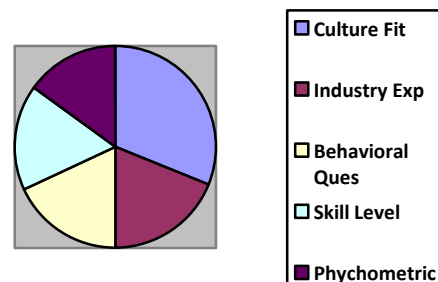

Chart -1: Survey

2. CONCEPTS RELATED TO CHATBOTS AND MACHINE INTELLIGENCE

Lately, Chatbots have ushered in a new era in the field of Machine Intelligence. For entitling the potential of an intelligent system, the system has the characteristics called the 'Performance Factor' which is a crucial and a highly important measure. Chatbots generally exhibit Perception of the data, Heuristic values, Memory, Consciousness in the surroundings. Chatbots are known as "Partially Intelligent Systems" because of the above stated reasons. Conventionally, an intelligent system can exhibit six fundamental traits of Intelligence according to the machine intelligence concept:

Comparison between the strings

Conversational Flow

Conversational Flow

Heuristic values

Memory within interactions

Consciousness

The equation below expresses that performance factor which is defined as a score given by the number of fundamental traits shown by the given system, divided by the total number fundamental traits of intelligence, i.e. Performance factor= number of fundamental traits (equation:1) displayed the values at the right-hand side of the equation can differ or vary from 0 to 1. According to the concept mentioned above, a "completely intelligent system" will have a performance factor of 1. According to our project, the chatbot can respond to 86431 out of 10,000 queries requested to it. This implies that the accuracy rate is 86.4.

3. LITERATURE SURVEY

In August 2015, Bogdan Ionescu et al. [1] essentially created conversational entity's architecture for synergetic rich text, media drafting, and for absorbing. 'THE CHAT' is the important keyword which plays an indispensable role to reconcile the group activities. The main objective is the development of information or content by one or all of the members of a group in a contemporize manner. The objects are exclusively web-based and contain information concerning partaking functions to facilitate the users in their synergistic work as well as associated with the embedding of objects in the content. In the paper [2] Haoliang Wang et.al has also planned for a conversational application called as "Touch Talk" which is an interaction design of a social TV. The main intention of touch talk is to provide comment sharing with each other in real-time while watching TV as well as encounter a better communication program. After assaying viewers behavior, some of the properties of social TV is examined. In accordance with the results of these research reasonable and meaningful platform exists of some principal designs for social TV. In the paper [3], G. B. Satrya et al. has confabulated about residual data from private chat, secret chat and the hidden chat in social messenger applications for Android and further gives an exegesis about the artifacts which are commenced by social media. Information about the generated messages, how they are analogous to one another and further provides interpretations among them.

4. RELATED THEORY
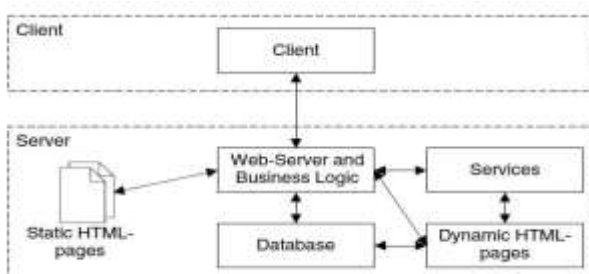
4.1 Web Server

Fig -1: Web Server Architecture

In a computer system, a Web server [4] is the entity that undertakes the requests through HTTP (Hypertext Transfer Protocol). For distributing the data or texts on world wide web (www), the web server proves to be an essential key. Web Server architecture is defined as a logical layout or design with the help of which a web server is designed, developed and deployed.

4.2 Database

A database is known as the warehouse of the data which is deployed in a structured and systematic manner which results in the data being easily retrieved, administered and streamlined. The data in the database can be updated, expanded, or deleted for the addition the new information. For the organization of data in a systematic manner, rows, columns, tables etc. are constructed. Indices are given to one or more constituting columns to make it easier to find relevant information. As new information is added data gets updated, expanded and deleted. A workload is given to the database processes for creating and updating themselves. It is also beneficial for querying the data they contain as well as in running applications against it.

4.3 Stack Overflow API

Stack Overflow [5] consists of millions of data entries corresponding to the questions asked by the user. It also comprises the responses provided as answers by other users. If the question asked is same as the one asked by previous users, an immediate answer can be fetched from the database. It is highly useful for the passionate programmers to deal with the real-world problems and find solutions to them efficiently.

4.4 Sentiment Analysis

To deduce whether the candidate's attitude is positive, negative, or neutral, we make use of sentiment analysis. To recognize and distinguish the responses of the candidate algorithmically in one word, we make use of sentiment analysis.

4.5 String matching

For weighing the resemblance among two entities, we make use of Dice coefficient [6]. The formula for the same is defined as twice the number of terms common to compared entities divided by the total number of terms in both tested entities. The value varies from 0 to 1. If the result is 1, it specifies identical vectors whereas 0 denotes vectors which are orthogonal or orthogonal vectors. Hence, Dice coefficient can be formulated as equation (2):

Dice coefficient = (2*Common Terms) / (Number of terms in String1 + Number of terms in String2)

$$\frac{2|X \cap Y|}{|X| + |Y|}$$

4.6 Software as a Service(SaaS)

Software as a Service is a Scalable and Efficient software distribution system. A third-party provider hosts the Software on the Cloud and makes it available to their customers over the Internet. This removes the need for organizations to set up, for running, maintaining and updating software themselves. This removes the high cost combined with the Pay per Use model where the organization only pay when they use the software and not otherwise.

4.7 Psychometric Score

These type of test papers are targeted for the graduates and job applicants. They check the mental ability of the candidate in terms of Numerical Reasoning, Verbal Reasoning, Diagrammatic Reasoning and Situational Judgement.

Numerical Reasoning: This includes one's ability to deal with numbers. It checks how accurate and quickly the candidate is able to handle the numbers. It includes various domains like currency, Ratio and proportion, Mixtures and allegation, percentages etc. this section contains 40 marks.

Verbal Reasoning: From this section, a firm can judge how constructive one can be taking the verbal section. This doesn't check one's ability of fluency or verbal communication. This includes reading a passage and then answering questions according to the paragraph. This section contains 20 marks

Diagrammatic Reasoning: To check or assess the logical reasoning capability of the candidate these are framed. In this one has to infer a set of rules from the flowchart given in the question and them constructively apply these to a new situation being asked. This section contains 20 marks Situational Judgement: They assess how the candidate will handle the situations at their workplace. Our test experts have identified 8 key competencies which are essential to get ahead of the competition. By this, a firm can decide whether the candidate is a perfect fit for the company. This section contains 40 marks.

The test is set for max score of 100 which is the psychometric score which is then used in the mathematical formula to find out the final output of the software.

5. Conceptual Framework
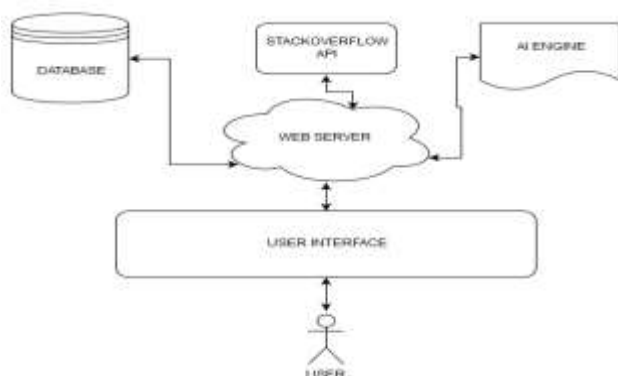
5.1 INTERVIEW CHATBOT ARCHITECTURE

Fig -2: Interview Chatbot Architecture

USER

The candidate who wishes to have a job at a particular firm is referred as User.

USER INTERFACE

This is the Chat interface where the user interacts with the chatbot, the chatbot asks questions and the user responds to those questions in a seamless manner.

WEB SERVER

Web Server provides the questions from the Database and displays it to the user on their screen through the user interface. When the user responds to the question, the response along with sentiment analysis is stored in the database.

AI ENGINE

Keyword Extraction of the messages sent by the users. String Matching, sentiment analysis, and grammar checking are done by the AI engine.

DATABASE

It is a collection of questions to be asked to the interviewee and stores all the answers given by the interviewee along with the keywords, grammar score, and sentiment score.

STACK OVERFLOW API

Questions which are asked in the interview are taken from stack overflow API, each answer contains 5 or more answers. The answers will be matched with user's answers with the help of string matching API.
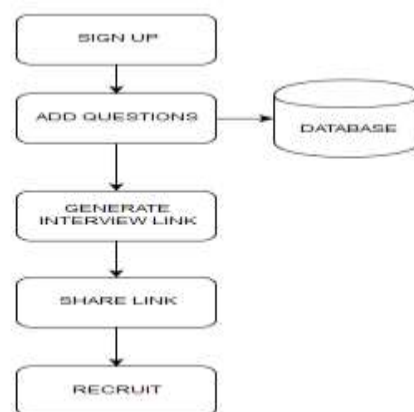
5.2 SAAS ARCHITECTURE:

Fig -3: SaaS Architecture

RECRUITER SIGNUP

The Interviewer/Recruiter creates an account on the SaaS Platform through a sign-up page. All the basic information about the candidate is asked in the sign-up process to maintain the data and information about the candidate and to ensure the authenticity of the employee to be recruited.

DASHBOARD

Recruiters can add questions which are to be asked to the interviewee. The recruiters can add as many questions as they want and desire to be asked from the candidate to check their skills and choose a deserving employee for their firm.

DATABASE

Consists of a collection of questions which are responded by the CANDIDATES. THE data in the database can be deleted, updated or expanded for the new information to be ADDED. THE database is also beneficial for querying the data.

INTERVIEW URL

A URL is created which can be shared by the recruiter so that the interviewing process can be handled in an efficient manner. An interview can be the part of the recruitment process at any point of time from any place.

RECRUIT

Recruiters can analyze and read the inputs given by the candidates, based on that recruiters can contact them for follow-up.

6. IMPLEMENTATION

As shown in Figure 3, The recruiter can create his account on the SaaS Platform, and add some questions needed to ask the candidates. All the questions are stored in the database and a link is generated which contains the AI that will conduct the interviewing process.

As soon as the candidate lands on the interview Artificial Intelligence chatbot link, he is greeted with welcome text and the interview begins.

The Artificial Intelligence asks questions and saves its answers, if there are more questions in the database even those are asked from the candidate. When no more questions are left the conversation ends and the recruiter can see all the details of the Interview on their SaaS Dashboard.

Now, when we talk in terms of candidate's point of view. After creating the account successfully, general information about the candidate is asked then comes technical and non-technical questions. To judge the candidate's responses and to know about the behavior of the candidate, Semantic Analysis is used. As shown in Figure 4, suppose the question asked is: Q1. What are your strengths and weaknesses?

Now the recruiter should get the response as positive, negative, or neutral regarding the candidate's response. The process followed for the above criteria is: As the candidate inputs his or her response, before being stored in database the responses reach the server. At the server, Semantic analysis is performed. The keywords are extracted and score is given according. The score of each word has been listed in AFINN, the summation of the scores is performed as displayed to us. If the score is greater than 0, it is considered as a positive response, similarly less than 0 indicated negative response and 0 denotes neutral. This is how a candidate is judged for a non-technical question or Sentiment Analysis is done for judging the candidate's behavior for non-technical

questions. The recruiter can know that the responses by the candidate are positive, negative or neutral.

Here in question 1, the score is -2 that means negative response whereas in question 2, the score generated after adding the scores of all the keywords is 8 which is highly positive, hence a positive response can be assumed.
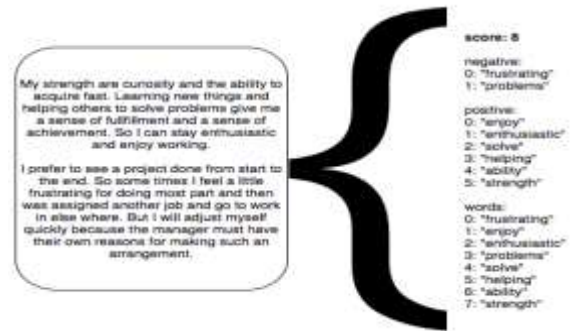

Fig -4: Output of the Question 2 using Sentiment Analysis.

The second most important part is the handling of technical questions. The responses given by the candidate are judged by matching the strings with the correct answers stored in the database. This is done by Dice coefficient, each word in the sentence is compared with the right answer. Even if the matching of one string gives the score greater than 0.6 than the answer is correct else it is a wrong answer. The Figure 6, illustrates the functioning of dice coefficient in the form of a flowchart. After getting the string from the user interface, the answer is matched with the answer in the database at the server itself. Then the condition comes whether it is greater than 0.6 i.e. 60%. No matter if the answer is right or wrong the responses are stored in the database. Accordingly, the firm can easily decide whether the candidate is a good fit for the firm or not.
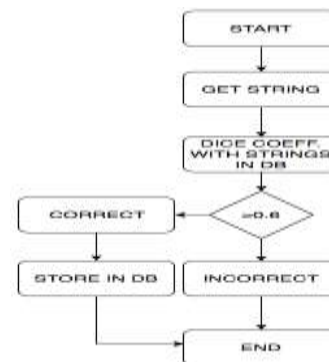

Fig -5: Functioning of Dice coefficient

Figure 7, illustrates the simplified version of working of the interviewing Chabot. This generic view is for the easier understanding of the users. After greetings, according to the responses given by the user the bot understands whether to ask further questions or not. If more questions are to be asked it goes back to ask question criteria else stops. For example, if the firm is looking for a software employee who codes in JAVA, if the question asked is: Do you know JAVA? If yes programming questions are asked else the interviewing

process is terminated. In this way the firm can end up getting a skilled employee to reach greater heights in future.
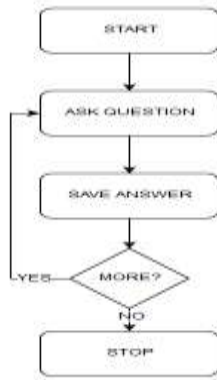

Fig -6: Interview Flowchart

As we found in our research, companies look for these criteria in their candidates in this order Culture Fit, Previous Industry Experience, Behavioral Questions, Skill Level and Psychometric Score. We use these five metrics to calculate the candidate recruit-ability score. Calculating the Weighted Mean of their individual values.

$$\frac{0.3CF + 0.2(BQ + IE) + 0.15(SK + PT)}{100}$$

Where, CF is score in Culture Fit, BQ is score in Behavioral Questions, IE is score in Industry Experience, SK is score in Skill Level and PS is score in Psychometric Score. All of these 5 scores are between 0-100.

This formula gives us a candidate recruit-ability score which ranges from 0-100. If a candidate has a score of >0.7, they are perfect for that Job

7. Result

Screenshot 1

The SaaS architecture with a sign-up page where the recruiter can sign up to create a chatbot for the interview was successfully created. The Link generated with the Artificial Intelligence chatbot was able to take interviews of the interviewee, the chatbot was able to fetch questions from the stack overflow API according to the knowledge of the Candidate. All the Answers were saved in a Database that could be accessed by the recruiter later.

8. Conclusion

With the emerging trend of online chatting which is becoming a crucial part of every living being and gaining popularity, where the customer can chat with the customer care representatives at any point of time from any place. This gives the freedom to the user to showcase their talent in a particular domain and thereby gaining job opportunities sitting at their location in a more efficient way. The proposed system architecture in this paper focuses on analyzing and implementing the interviewing chatbot by identifying whether the skills mentioned by the customers are rightly identified or not. All the responses from the candidate are sent to the Chatbot which try to identify whether the candidate can prove to be a suitable employee to an organization by having the conversation with the candidate and judging them on both technical and non-technical perspective. This, in turn, saves a lot of money and resources of the company used for identifying an efficient candidate for an organization and even making the customer satisfied with their job. As this proposed system is implemented on AWS[8], it makes this system capable of handling an enormous amount of user base.

REFERENCES

[1] Bogdan Ionescu, Cristian Gadea and Bogdan Solomon, IEEE 10th Jubilee International Symposium on Applied Computational Intelligence and Informatics (SACI) "A chat-centric collaborative environment for web-based real-time collaboration,", Timisoara, Romania, Aug 2015. Retrieved Nov 2017.

[2] Haoliang Wang, Chunhong Zhang and Ming Li, International Symposium on WPMC, " Social TV real-time chatting application design," Jan 2015. Retrieved Dec 2017.

[3] G. B. Satrya, P. T. Daely and S. Y. Shin, Eighth (ICUFN 2016)International Conference on Ubiquitous and Future Networks 2016, "Android forensics analysis: Private chat on social messenger", Vienna, Austria, 2016. Retrieved Dec 2017.

[4] https://www.webdevelopersnotes.com/what-is-web-server, "Overview on Web Servers", 2018. Retrieved Jan 2018.

[5] https://api.stackexchange.com/docs, "Overview of Stack Exchange API v2.2", 2018. Retrieved Jan 2018.

[6] http://nlp.fi.muni.cz/raslan/2008/raslan08.pdf#page=14, "RASLAN 2008, Recent Advances in Slavonic

[7] Natural Language Processing, Masaryk University, 2008, Second Workshop on Recent Advances in Slavonic, Natural Language Processing, RASLAN 2008 Karlova Studánka, Czech Republic, December 5–7, 2008 Proceedings. Retrieved Feb 2018.

[8] https://aws.amazon.com/documentation, "Documentation on Amazon Web Services", 2018. Retrieved Feb 2018