Automation in the Interview Assessment Process

Priyam Shah[1], Mihir Shah[2], Amit Bhujbal[3], Rohini Nair[4]

[1,2,3]Student, Dept. of Computer Engineering, K J Somaiya College Of Engineering, Mumbai, India
[4]Professor, Dept. of Computer Engineering, K J Somaiya College Of Engineering, Mumbai, India

---***---

Abstract - *Interviews consume the time and resources of the company. The automation of the pre-screening rounds can save the cost and time of the companies. Fundamentally, recruitment for a specific job role requires the candidate to know the hard and soft skills required for that job. These can be configured in the system and based on AI assessment algorithms filtration can be done for ahead rounds. We have proposed two algorithms for assessing the soft and hard skills of the candidates in this paper. The personality insights algorithm helps in assessing the OCEAN skills[3], and the answer relevancy algorithm assesses the significance of answered questions by a candidate. We implemented a website to simulate the recruitment process with the developed algorithms.*

Key Words: Recruitment, Information Retrieval, TF-IDF, Cosine Similarity, NLP, Video Interview, OCEAN personality traits.

1.Introduction

We all know that hiring someone costs money. Recruiters must post job openings on job boards, conduct interviews, do background checks, and, of course, distribute and recover all paperwork related to hiring a new employee. According to Bersin by Deloitte, the average cost per hire is about $4,000. Depending on the job level and hiring processes, the number will vary, but every hiring manager may relate to the problem of inflated talent acquisition costs. Why does hiring a new employee, which is supposed to be a cost-cutting measure, consume so much time and money? Why do accountants and financial executives shudder at the prospect of conducting a search for a new employee?[1]

The maximum time is consumed in the background and knowledge examination to find the best fit for the job responsibilities and company. In the pre-screening interviews, the interviewer has to assess 50-100 candidates to verify their soft and hard skills. We aim to automate this process by reducing the required time and manpower. We plan to achieve this by implementing an AI-based skill scorer to give insights of multiple candidates within single submission rather than taking multiple interviews. This will help the companies to manage resources (manpower and time) and save cost-effectively.

The paper is divided into four main sections. Sections two and three describe the background study of various products and algorithms in this field. The fourth and fifth section describes our implementation and its results.

1.1 Hard and Soft Skills

Soft skills, often known as interpersonal skills, enable employees to make the most of their hard skills. Skills like dispute resolution, emotional intelligence, time management, and functioning well under pressure are essential in the job. Candidates can use soft skill interview questions to demonstrate personality attributes that can be utilized in the workplace. Soft-skilled employees become terrific team players and efficient bosses. They are skilled at forming bonds and are easy to get along with. They are also fantastic brand advocates who have the opportunity to advance in their roles and at their working organization. Employers should focus on the seven important soft skills: communication skills, emotional intelligence, team player, growth mindset, time management, creativity, and leadership. [2]

Hard skills, also known as subjective skills, are quantifiable abilities or skill sets that may be taught. These abilities are typically obtained through classroom learning, reading, or on-the-job training and are frequently stated in a cover letter and CV to help employers discover them. Proficiency in a foreign language, Computer Programming, as well as a degree or certificate are examples of hard skills. [2]

These skills set the basic layout for assessing an ideal candidate for the required job.

2. Survey of available products

MyWays [4]: It is a recruitment portal for internships. Employers may advertise internships, review individuals who have applied for the internship, and hire them. This is an old way of recruiting in which just the internship information is published, and the rest of the interview process has to be performed manually by recruiters for every candidate. We recognized the need for automation in this process from this website. Autogram [5]: On this portal, recruiters can create general job posts, and candidates can apply for them by answering questions in a video interview. The recruiters then manually assess the interviews and add a comment for their team about candidate evaluation. Further, the evaluators collaboratively take a decision to hire/reject the candidate.

Video Interview [6]: A modern-day interactive solution for video interview-based hiring. The recruiter can grade on parameters like attitude, expertise, and spontaneity of the candidate's answer and collaboratively hire/reject a candidate with his team. Adaface [7]: This platform uses a chatbot-based UI for screening. Only coding and aptitude questions are asked and reviewed automatically based on answers preset during the question configuration step. Additionally, a library feature is provided where the sample questions and packages are available for direct conduction of general job profiles.

EasyHire.me [8]: Among other alternate recruitment websites, EasyHire focuses more on interview setup steps more in detail. During the assessment of candidates, the recruiter can send emails for hire/reject, set up an online interview, schedule an onsite interview and many more feasible options. Reculta [9]: It aims at helping institutions and companies come together to conduct hassle-free recruitment processes. A cloud-based solution that helps boost efficiencies in placement processes through automation and organizes all placement-related activities using smart dashboards and advanced technology.

By surveying all these products, we extracted some distinct and innovative features from each product that would help us develop a definite solution for automating the recruitment process. We needed a Video Interview based assessment method over the manual interview process, a sample question library for multiple job profiles, collaborative feedback on the assessment by team members, and simple cloud-based solutions that can be easily deployed and used from anywhere in the world.

3. Literature Survey

This section focuses on research papers studied for Question-Answer Relevancy Algorithm and Personality Assessment Algorithm.

IBM Watson Personality Insights for OCEAN Skills. [3] This study gives an in-depth view of the OCEAN personality Insights of a human based on his language. The frequency with which people use certain categories of words can provide clues to their characteristics. The Personality Insights API analyzes the content, and based on the pre-trained models, the IBM gives the personality insights scores for OCEAN skills. In the paper written by Hrazdil [12], they measure the OCEAN traits for a large sample of CEOs and CFOs, to compute a measure for risk-tolerance. They conducted a number of validation tests to show that their risk-tolerance measure varies with the existing inherent and behavioral-based measures in predictable ways; that all of the firm-year level personality trait measures are manager-specific and are not related to firm characteristics and firm performance; and that their estimate of the risk-tolerance trait supports the association between CEO risk tolerance and audit fees. Specifically, they found that higher CEO risk tolerance is associated with significantly higher audit fees and that CEO personality has an incremental impact on auditors' assessment of client risk beyond the risk-taking incentives induced by their compensation portfolios.

The authors of the paper [13] proposed a solution that uses an intelligent chatbot that drives the screening interview. The users (job candidates) will feel like they talk to a real person and not just fill a simple web form for another job interview. At the same time, the chatbot can evaluate the data provided by users and score them through a sentiment analysis algorithm based on the IBM Watson Personality Insights service. Their solution replaces the first step in the interviewing process and automatically elaborates a candidate profile.

A Naïve Approach for Monolingual Question Answering. [10] This paper follows the approach in which, first of all, the question is categorized into Factoid, Definitive, Reason, Procedure, or Purpose, then its keywords are fed to Lucene Indexer to find the relevant passages from an already indexed database of questions and answers. Passages are then ranked based on the density of keywords matched, and the most appropriate passage is used to check answer similarity. A more advanced algorithm over previous uses NLP to score answers [11]. It uses an NLP-based algorithm to find answers for user queries from the knowledge base. Here, the knowledge base is generally defined as Wikipedia, Microsoft Satori, Freebase. It provides an efficient scoring method to calculate the average score for the sentences from the answers found in the knowledge base.

The authors of the paper [14] summarized 130 research studies on the current state of the Question Answering literature. The observation gave insights that the natural language approach is the most flexible approach to independent language implementation algorithms. Besides the paradigms and their approaches, they were able to enlighten the metrics used to evaluate QA systems, the domains in which researchers apply their systems, and how NLP is used in this task, apart from the paradigm adopted.

Based on the study, we concluded that for the personality insights algorithm, OCEAN traits must be considered a key parameter for results, and an NLP-based algorithm must be used for answer relevancy scoring.

4. Practical Application

We have created a website and algorithms to simulate the complete recruitment process in our current implementation. For the screening rounds, the recruiter can create questions to test the hard and soft skills of candidates using the question banks. The candidate has to answer those questions in a given timeframe during recording. We then apply our assessment algorithms on recorded video interviews: personality test algorithm for soft skills assessment & answer relevancy algorithm for hard skills assessment. The overall scores generated for skills will help the recruiter to filter out the best suitable candidate for the job.

4.1 Architecture

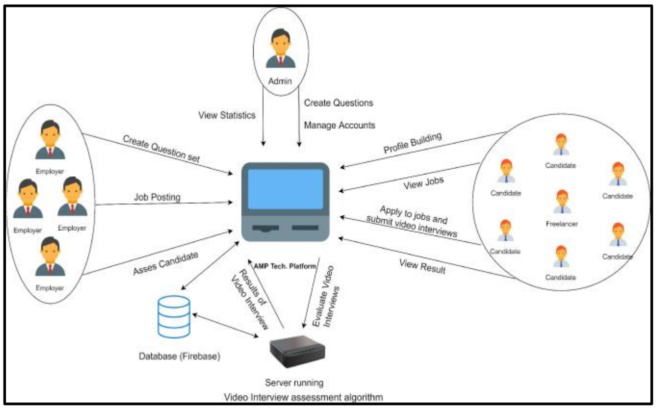

Fig 1: AMP Tech application architecture

We have named our system as AMP Tech and there are 4 major modules in our system:

1. Recruiter Module
2. Candidate Module
3. Admin Module

Recruiter Module - Figure 2 shows the functionality available with the recruiter persona to simulate the entire recruitment process on our system. The recruiter can create a job opening post, create a question set for the interview, and then based on the submissions of candidates and their results, he/she can filter out the candidates for that job post.

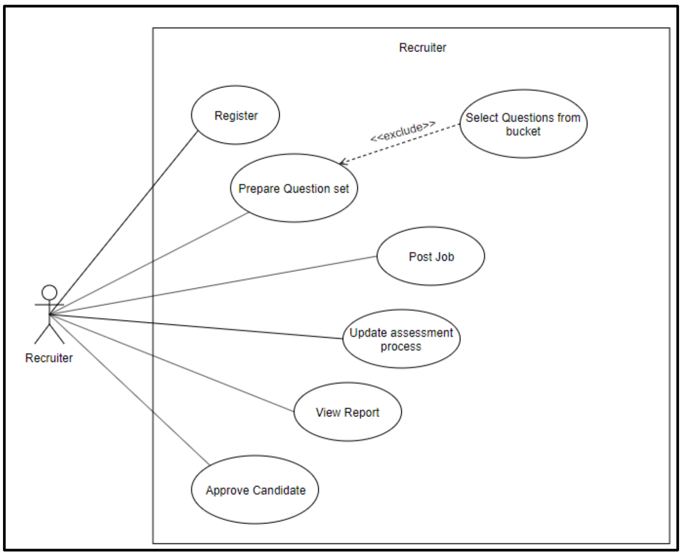

Fig 2: Use case diagram for Recruiter Persona of the application

Candidate Module - Figure 3 shows the functionality available with the candidate persona. The candidate will be able to create his profile/resume and then apply for the job openings based on his profile. The candidate then needs to answer a video interview for the given job profile, which will be processed by our system, and the scores will be shown to the recruiter.

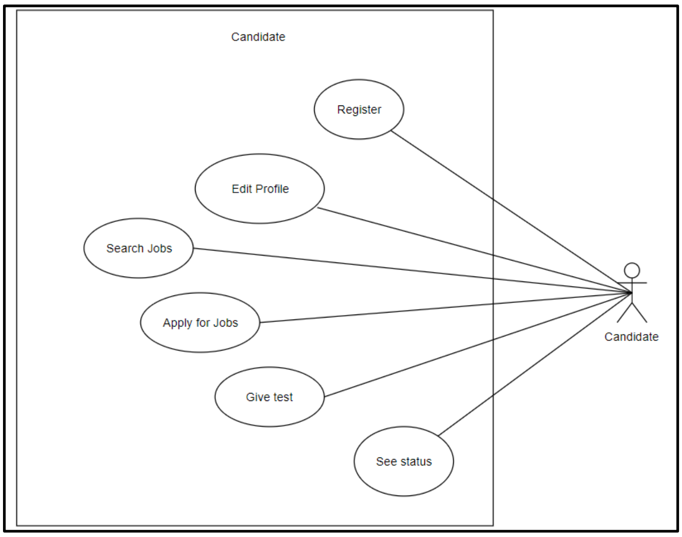

Fig 3: Use case diagram for Candidate Persona of the Application

Admin Module - Figure 4 shows the functionality available with the admin persona. Admin will get an analytical overview of the entire system. Admin can manage the user accounts block/unblock in case of discrepancy. Admin can create or update the library of questions for the general job profiles, which the recruiters will use as a base questionnaire to generate a test for the job post..

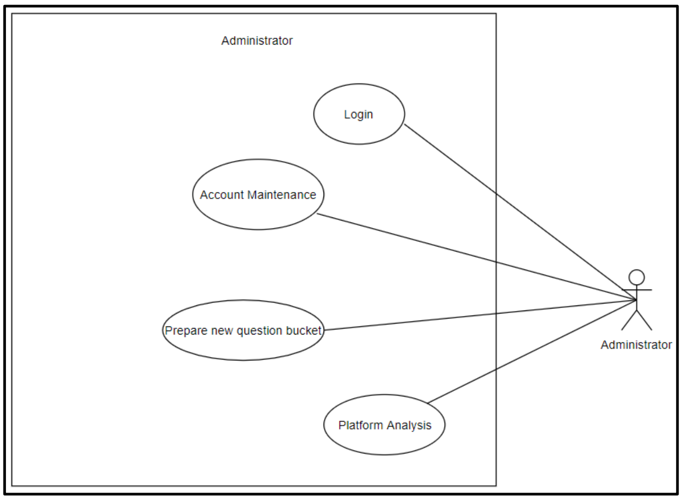

Fig 4: Use case diagram for Admin Persona of the Application

4.2 Algorithms

We have proposed two AI-based assessment algorithms for skill scoring. First is the Answer Relevancy Algorithm for scoring hard skills on a scale of 5 points. The second is the Personality Insights Algorithm to score the soft skills (OCEAN values) on a scale of 5. The algorithms were implemented in Python 3.

4.2.1 Answer Relevancy Algorithm

This algorithm in our system assesses the hard skill questions labeled by the recruiter. The video-interview answers for such questions will be passed as an input to the algorithm where initially pre-processing will be done like converting video to audio and then audio to text. The question will be passed to the Search API and based on various answers fetched, the keywords would be extracted from those answers, and a bag of stemmed words would be created. This bag of words will be used for relevancy matching on answered video text. Later on, we plan to add additional weights for keywords to improve the accuracy.

Figure 5 explains part 1 of the algorithm for data preparation. It can be explained as following:

Search question: The question is searched on google using the google search query by the algorithm, and first page URLs are extracted from the result of the search query. These URLs are stored in an array and passed to the summarizer one by one.

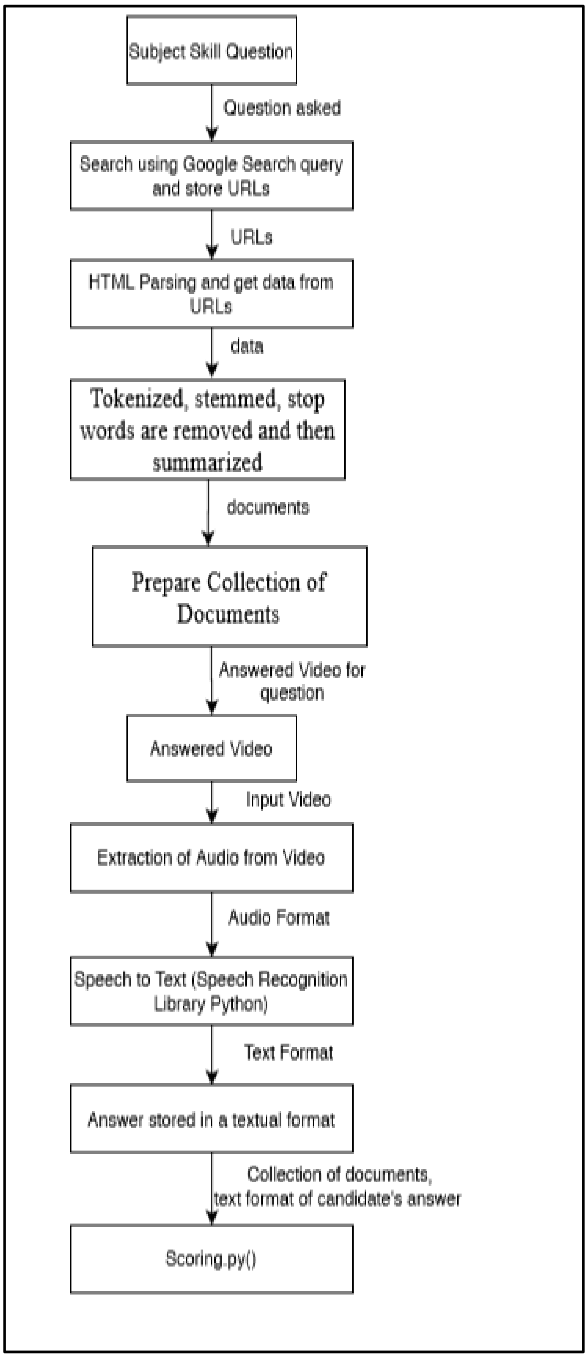

Fig 5: Flowchart for Part 1 of Answer Relevancy Algorithm (Data-Preparation)

Summarize from URL: Each URL is parsed using HTMLParser and retrieves its data. These sentences are tokenized, stemmed, and stop words are removed and then summarized. The summarized data is stored, and finally, a collection of such documents of relevant answers is prepared. Now we will convert the video answer uploaded by the candidate to text.

Video to Text conversion: Reading clip from the video using VideoFileClip function from the moviePy library. This clip is used to extract the audio and write it in an audio file saved in wav format. Using the speech recognition library provided by Google, the audio file is read and saved in textual format.

After this step, the answer in the textual format is preprocessed and compared with the collection of documents of relevant answers to generate a score. Collection of documents of relevant answers and candidate's answer in the textual format is passed for scoring algorithm in part 2 as input.

Figure 6 explains the flowchart for the part 2 of the answer relevancy algorithm which is scoring. It can be explained as following:

Processing of relevant answers: There are various steps for pre-processing of documents in a collection.

a) convert_lower_case(): conversion of complement document in a uniform case.

b) remove_punctuation(): punctuation signs in the document are removed using regex.

c) remove_apostrophe(): removal of apostrophe from the document.

d) remove_stop_words(): Remove stop words in English from the document.

e) convert_numbers(): Conversion of numbers to words for eg.100: Hundred

f) stemming(): stemming of words.

After this, removal of punctuation, conversion of numbers, stemming, and stop word removal are needed again. Because some of the previous methods can again create the punctuation or stop words e.g., 121: One hundred and twenty-one, in this case, punctuation (- hyphen) and stop word (and) are added.

The result of all the above methods create a tokenized preprocessed word document, and a collection of tokenized documents of relevant answers is created.

Calculating DF: Iterate through all the words in all the documents and store the document ids for each word. A dictionary can use the word as the key and set of documents as the value. The set will not just take duplicate values, even if the addition of the document takes place multiple times.

Calculating TF-IDF: In information retrieval, TF-IDF, short for term frequency-inverse document frequency, is a numerical statistic that is intended to reflect how important a word is to a document in a collection.

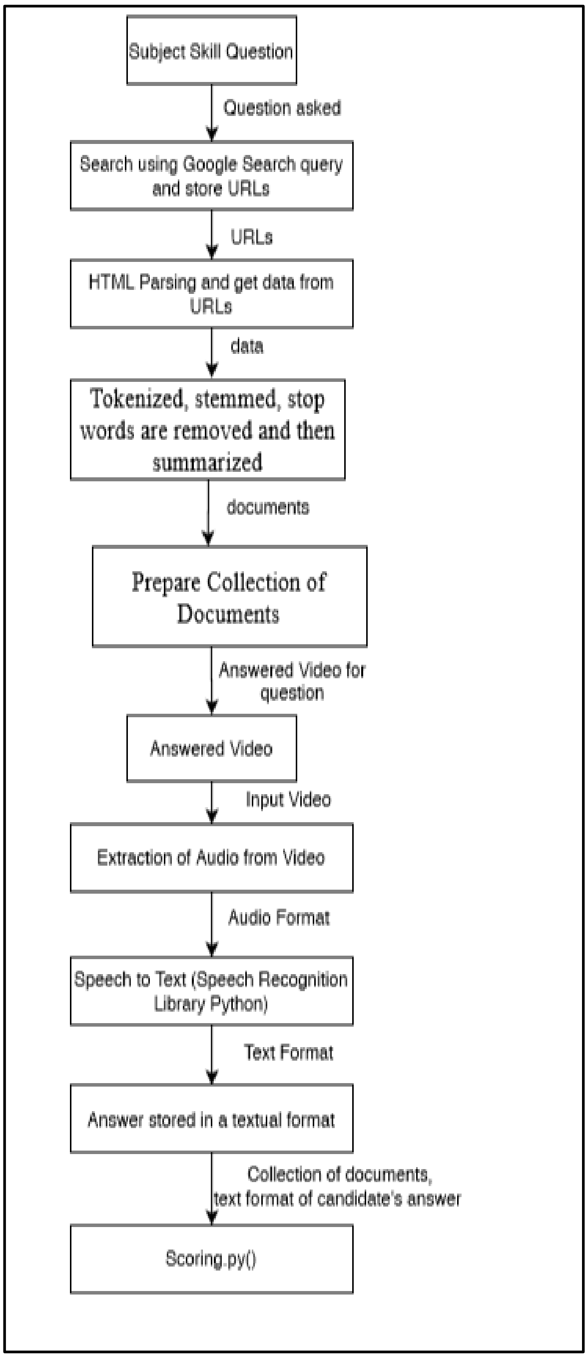

Fig 6: Flowchart for Part 2 of answer relevancy algorithm (scoring)

To calculate TF-IDF, the dictionary is used with (document, candidate_answer) pair as key and any TF-IDF score as the value. We just need to iterate over all the documents, and we can use the counter which can give us the frequency of the tokens, calculate TF and IDF and finally store them as a (document, candidate_answer) pair in TF-IDF.

Cosine Similarity: Cosine similarity is used to mark all the documents as vectors of TF-IDF tokens and to plot them from the center of origin as shown in Figure 7. Sometimes, the query length would be small but it might be closely related to the document, in these cases, cosine similarity is the best to find relevance.

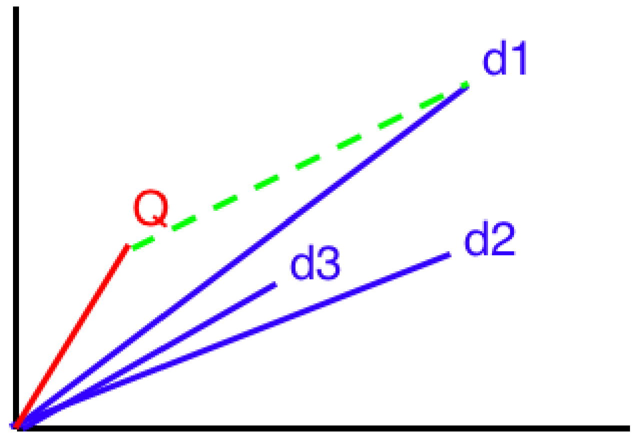

Fig 7: Representation of Cosine Similarity of vectors

a) Vectorization: To compute the cosine similarity, the simplest way is to convert all the documents in a vectorized format. The TF-IDF values of the collection of documents are used to find the vector and stored in a NumPy array format.

b) Pre-processing of relevant answers: Before vectorizing the candidate's answer, it should be in the same format as the collection of documents. Hence, we need to pre-process the answer text

c) Generate a vector for the candidate's answers: Using the method of vectorization a vector for the candidate's answer is to be generated to measure the angle between the vector of answer and vectors of documents.

d) Now, we have to calculate the cosine similarity for all the documents with the query. Cosine similarity is defined as follows. np.dot(a, b)/(norm(a)*norm(b))

e) It measures the cosine of the angle between two vectors projected in a multi-dimensional space.

f) The cosine of the angle between the vectors is stored in an array. Two documents are similar if the angle between them is less.

g) The lesser the angle, the higher is the cosine value. Hence, the document id is retrieved having the maximum cosine value. The score is generated by normalizing the cosine value on a scale of 5.

The final average score of all hard skill questions answered by the candidate is then displayed to the recruiter on a scale of 5.

4.2.2 Personality Insights Algorithm

The soft skill question assigned by the recruiter to assess the candidate for the job via video interview is mapped to be assessed by this algorithm in our system. The answered video of the candidate for the soft skill-specific question will be passed as an input to the algorithm where initially pre-processing will be done like converting video to audio then audio to text and then passing that text to IBM Watson API for getting JSON response as the output of OCEAN values in percentage. Figure 8 shows the flowchart for the implementation of the algorithm.

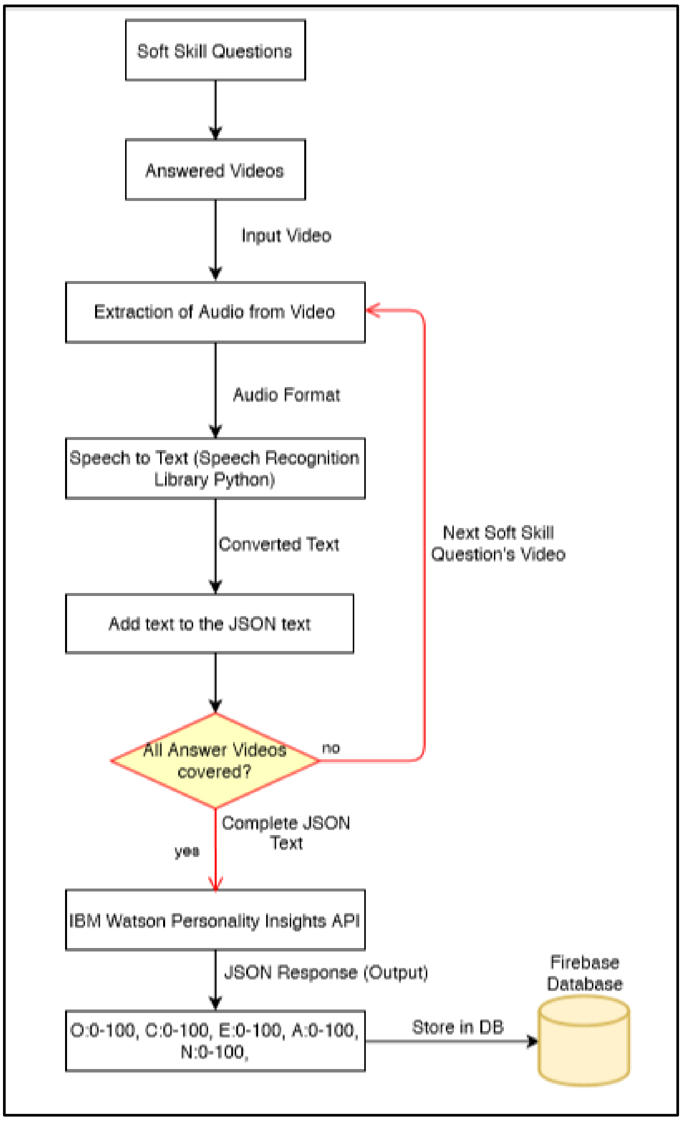

Fig 8: Flowchart for Personality Insights Algorithm

The video answers by the candidate for the soft skill question are processed using IBM Watson Personality Insights API to get then OCEAN values. Each value is between 0-100, which later is normalized to a scale of 5.

O: Openness, C: Conscientiousness, E: Extroversion, A: Agreeableness, N: Neuroticism

Pseudocode for Personality Insights Algorithm

a) Reading clip from the video using VideoFileClip function from the moviepy library. This step extracts the audio and writes it in an audio file saved in WAV format.

b) Using the speech recognition library provided by Google, the audio file is read and saved in textual format. After converting all the soft skill questions to a textual format, the complete text file is saved in a JSON format.

c) After this, the complete JSON text is sent to the API for the generation of the OCEAN values.

d) IBM Watson Personality is registered using the API key, and JSON text is passed to the API.

e) The API returns the OCEAN values of personality traits in JSON format.

f) Each value is extracted, and the percentage is calculated.

g) The scores are normalized and shown to the recruiter.

5. Results

The practical implementation of the system helped in testing the algorithms for five job profiles Software Developer, Accountant, Business Consultant, Teaching Professional, and Event Manager, over 10 candidates each. We had limited resources to validate our work, so we performed the manual assessments with experts in the field as a benchmark to compete with our algorithm-generated results. The accuracy was 79% for finding the best-fit candidates using these algorithms with properly configured skill-sets. It is a decent accuracy for algorithms running on an open dataset. We expect to test it for broader job profiles and with thousands of candidates to get more accurate efficiency of the algorithm.

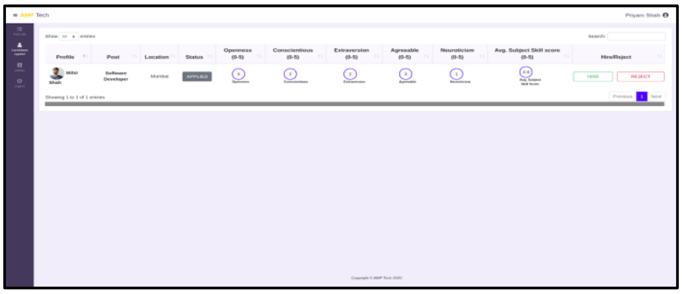

Fig 9: Display of performance results to the recruiter for a specific job profile posted.

Figure 9 shows the display screen for the performance results shown to the recruiter for their posted jobs. It shows the list of candidates who gave the interview with filters and sorts for the recruiter to select the best fit for the role or subsequent rounds.

6. Conclusions

The project focuses on one of the most prominent parts of the human resource industry, which is recruitment. The project automates the time-consuming part of the recruitment process except leaving the crucial parts of the process like posting job requirements and selecting the final candidate still in the recruiter's hands.

A conventional interview process of 10 candidates with 1 recruiter could take 5 hours optimally if allocated 30 mins per candidate without a break, whereas our system brings that time to 40 mins to evaluate all candidates at once. We have built a website and algorithms to demonstrate the complete use case. The solution drastically reduces manpower and time duration in the recruitment process by providing an AI-based analysis of each candidate’s application for the job. The Personality Insights Algorithm provides a BIG 5 personality values for the soft skill-based question. The Answer Relevancy Algorithm provides a relevancy score for the subject knowledge-based questions. The scored values help the recruiter easily find the best-fit candidates to hire for the job, thus making the recruitment process easy for both candidates and recruiters.

7. Future Work

Since this project is a unique idea and a huge time & cost saver for all people involved in the job recruitment process, the project has a large future scope. The project can be made more robust and secured. Collaborations can be done with a few companies to test its real-world performance and introduce it as a complete weapon for the recruitment process. Currently, the system has algorithms for processing the hard skill and soft skill questions. But later, we can even include the processing of logical and analytical questions, which will cover a broader aspect of the interview questions.

This system is made to help the recruiter majorly, but we can also make it a tool for educational purposes. The students (before entering the recruitment phase) can give mock interviews on the system and can assess themselves based on the score for their answers, and the candidate can then practice the points or skills which he/she is lacking and give the interview again and try to become perfect at it. This will give them confidence for the real interviews. Colleges can use this educational tool to train their students.

Face Recognition feature can be added to recognize and match the candidate's photo submitted in the resume, and the person giving the video interview is the same. This will not permit candidates to cheat the system and recruiter both. Thus, a couple of future works are mentioned above, but the HR industry is enormous, and there can be tons of improvisation made for this product.

REFERENCES

[1] Glassdoor for Employers, "The Hidden Costs of Onboarding a New Employee", accessed 26 December 2021, https://www.glassdoor.com/employers/blog/hidden-costs-employee-onboarding-reduce/

[2] The balance careers, "Hard Skills vs. Soft Skills: What's the Difference?", accessed 26 December 2021, https://www.thebalancecareers.com/hard-skills-vs-soft-skills-2063780

[3] IBM Watson Personality Insights document, 2019-07-30:https://cloud.ibm.com/docs/services/personality-insights?topic=personality-insights-science

[4] MyWays, accessed 20 April 2019, myways.in

[5] Autogram, accessed 23 April 2019, myautogram.com

[6] Video Interview, accessed 23 April 2019, videointerview.io

[7] Adaface, accessed 25 April 2019, adaface.com

[8] Easyhire.me, accessed 3 May 2019, easyhire.me

[9] Reculta, accessed 5 May 2019, reculta.com

[10] Sanglap Sarkar, Venkateshwar Rao Madasu, Baala Mithra SM, Subrahmanya VRK Rao. NLP Algorithm Based Question and Answering System. Seventh International Conference on Computational Intelligence, Modelling and Simulation, 2015.

[11] Rohit Bharadwaj, Surya Ganesh, Vasudeva Varma. A Naïve Approach for Monolingual Question Answering. Working Notes for the CLEF, 2009.

[12] Hrazdil, Karel, Jiri Novak, Rafael Rogo, Christine Wiedman, and Ray Zhang. "Measuring executive personality using machine-learning algorithms: A new approach and audit fee-based validation tests." Journal of Business Finance & Accounting 47, no. 3-4 (2020): 519-544.

[13] Carțiș, Andrei-Ionuț, and Dan Mircea Suciu. "Chatbots as a Job Candidate Evaluation Tool." In OTM Confederated International Conferences" On the Move to Meaningful Internet Systems", pp. 189-193. Springer, Cham, 2019.

[14] Soares, Marco Antonio Calijorne, and Fernando Silva Parreiras. "A literature review on question answering techniques, paradigms and systems." Journal of King Saud University-Computer and Information Sciences 32, no. 6 (2020): 635-646.