IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

HIRE MATE

Pioneering the Future of Talent Acquisition with AI-Powered Solutions

[1]Abijith S, [2]Adithya, [3]Adithyan V, [4]Anirudhan P, [5]Vipin K M
[1]Student, [2]Student, [3]Student, [4]Student, [5]Assistant Professor (CSE)
[1]Computer Science and Engineering Department,
[1]Nehru College of Engineering and Research Centre (NCERC), Thrissur, India

Abstract: The fundamental prerequisites for today's hiring procedures for a specific job post are chatbots for interviews and resume screening. Many applicants submit their resumes, and it takes a lot of work to manually sift them according to the job requirements to narrow the field of candidates to choose from. The project's driving force is its ability to overcome the difficulty of automating the hiring process, from screening resumes to testing candidates on their first technical exam. The difficulty in vetting resumes would be in applying Natural Language Processing to extract structured data from a variety of resume formats and extract all the keywords related to programming languages. The creation of questions and the application of Natural Language Processing tools to assess the candidates' responses present the other issue. Finding the candidates' eye movements during the interview and lip-reading them to look for any unethical behavior is the difficult part of proctoring a test. The objective here is to extract concept terms from the resume and search all keywords related to programming languages from the applicant's resume. The candidate would then be assigned a work role that fits his or her skill set based on that, which would shorten the time needed to find the right individual for a particular job description. The purpose of the candidate interview would be to learn more about the individual's technical abilities. Tests with multiple-choice questions could be utilized for the preliminary screening. However, they are not as effective in evaluating the candidate's depth of knowledge as the short-response questions are. Thus, it's critical to create assessments and short-response questions. Either sentence-level or paragraph-level information would be the basis for the question that would be generated for the candidate; the sentences and paragraphs would be on various technical topics on programming languages and concepts. Both at random and in reaction to the applicants' answers, the questions would be chosen. The camera would be used to proctor the exam during the interview, ensuring that no more than one person was in front of the camera and that the candidate was not using any other devices. The candidate's ultimate score would be determined by how well the job description and candidate interview aligned.

Keywords: CNN (Convolutional Neural Networks), NER (named entity recognition), BERT (Bidirectional Encoder Representations from Transformers), Talent Acquisition, Hiring automation

I. INTRODUCTION

Given the current state of affairs, the sector needs to automate the hiring process more and more. For a certain job position, multiple applications are available at once. Companies invest a lot of money in the recruitment process to find the best candidates for open positions because it is impractical to manually categorize those resumes according to the skill requirements. They hire candidates through online recruitment or by assigning a contract to a third party. Several techniques use resume categorization to determine if an applicant is a good fit for the position. They start by addressing the issue of resumes being unstructured, which makes it difficult to extract the necessary data from them to determine a candidate's skill set and assign a job role [1]. This helps to reduce the amount of work that personnel have to do to handle the candidates. The first method of resume matching involved matching all potential openings with the abilities collected from resumes worldwide. However, the issue with this strategy is that it takes a lot of time to complete the work. The latter divides the resume into several categories according to skill sets, and it does the same for job openings depending on the necessary talents. Additionally, it is categorized, and only those categories are used for resume matching [1][2]. Matching time is reduced with this method. The percentage weights assigned to a resume's skill set and title will have a big influence on how well resumes match when determining a target score [1]. Text chatbots with natural language processing (NLP) capabilities can be used in the recruiting process to screen resumes, extract pertinent information, ask candidates questions depending on their talents, and provide a final score [3].

II. LITERATURE SURVEY

A literature review is an overview of the works that recognized academics and researchers have published on a certain subject. It comprises the state of the art, encompassing significant discoveries as well as theoretical and methodological advancements on a given subject. Reviews of the literature do not present newly conducted experiments; instead, they rely on secondary sources. A literature review enables us to improve and showcase our abilities in two primary domains: locating knowledge and evaluating it critically.

2.1 JRC: A JOB POST AND RESUME CLASSIFICATION SYSTEM FOR ONLINE RECRUITMENT

[1]The overwhelming number of unstructured resumes that are flooding job sites presents issues for the current online recruitment landscape. Since candidates from a wide range of specializations submit resumes in a variety of forms and styles, it is challenging for traditional approaches to effectively extract and arrange pertinent data. This paper proposes the JRC (Job Post and Resume Classification System) as a solution to this problem. JRC uses an integrated knowledge base to make the categorization procedure go more quickly. JRC matches resumes to their respective occupational groups; in contrast to traditional systems, which scan the full range of resumes and job postings. With the use of automation in the screening process and resume categorization, this focused strategy seeks to reduce the workload for employers.

The study describes tests carried out with actual recruitment data to demonstrate the effectiveness of the suggested JRC method. These tests assess the system's efficacy and efficiency in comparison to existing online recruitment platforms, showcasing its benefits for candidate classification, screening, and matching to job ads in related occupational domains.

2.2 A MACHINE LEARNING APPROACH FOR AUTOMATION OF RESUME RECOMMENDATION SYSTEM

[2] Selecting the ideal candidate from a large pool of candidates for a vacant position is frequently a difficult task that impedes team development. But there's hope that an automated "Resume Classification and Matching" system can expedite this laborious and time-consuming procedure. This system uses a variety of classifiers to effectively sort a huge number of resumes into relevant job categories, which helps with the screening and shortlisting process.

The system uses content-based recommendation algorithms to further narrow the selection after the initial categorization is finished. This is utilizing a metric called cosine similarity, which evaluates how similar resume material is to job descriptions. The technology ranks the candidates according to this similarity, making it possible to identify the best matches depending on how well they meet the job requirements. Additionally, using k-NN (k-Nearest Neighbors) makes it easier to find resumes that closely match the job description that is given. This sophisticated method presents the best candidates for consideration, which speeds up the selection of applicants and aids in the decision-making process.

This automated system serves two purposes: first, it uses a variety of classifiers to divide up resumes into segments linked to different jobs; second, it uses content-based recommendation algorithms to rank and select the best candidates who closely match the job description. This methodology increases the chances of finding the best candidates for the available position while also speeding up the candidate selection process.

2.3 INTERACTIVE INTERVIEW CHATBOT

[3] The interview process is an essential means of vetting applicants for open positions and providing a first indication of their appropriateness. Traditionally, there have been two main components to this process: technical interviews, which concentrate on certain job-related skills, and human resource interviews, which assess a candidate's general fitness and fit into the business culture. Because this traditional strategy is frequently complicated, time-consuming, and resource-intensive, there is a need to investigate new simplified, and effective approaches. Using technology to automate the interview process is one such creative strategy. The suggested remedy entails using chatbots in place of human interviewers to give candidates a more uniform and structured engagement. The system attempts to provide an interactive interview experience by using Google Text-to-Speech and a chatbot with Natural Language Processing (NLP) capabilities. To replicate an actual interview setting, candidates converse with the chatbot and reply to questions and prompts conversationally.

In addition to performing the interview, the chatbot also creates thorough reports based on the responses from candidates. The recruitment staff can gain useful insights into candidate performance and suitability from these reports. The produced data helps with the unbiased evaluation of applicants, allowing for a more knowledgeable and data-driven shortlisting procedure. In the end, this system replaces human interviewers with a chatbot by utilizing the power of NLP and technology to create an interactive interview environment. This simplifies the interview procedure, creates consistent interactions, produces comprehensive results, and makes it easier to evaluate and select candidates more effectively and uniformly.

2.4 HR-BASED INTERACTIVE CHATBOT (POWERBOT)

[4] Artificial Intelligence, especially in Computer Science, has become a prominent field in the field of technical development, bringing with it several innovations. Because they mimic human-like user interactions, chatbots—a unique kind of computer application—have become more and more popular. With the help of verbose operator programs, these bots converse with consumers in a natural-sounding manner to make the contact seem seamless and as though the users are speaking with real people rather than machines. Carefully examining user input to make sure the bot responds in an appropriate and meaningful way is essential to generating this appearance of natural dialogue.

The goal of this project is to leverage artificial intelligence capabilities, namely natural language understanding techniques, to create an HR chatbot. Through the use of natural language inputs, consumers can engage with the chatbot to provide a tailored and responsive experience. The chatbot's function goes beyond simple communication; it is a comprehensive tool that allows users to conveniently manage a variety of work-related tasks and obtain extensive company information. Employees frequently run across obstacles when attempting to complete fundamental obligations, which calls for conversations with team leaders or HR staff. By allowing workers to do duties like requesting leave, compensation, allowances, and more directly through the chatbot, this chatbot aims to reduce such difficulties.

Additionally, the goal of this HR chatbot is to facilitate communication between staff members and HR representatives by offering a smooth avenue for support requests, query responses, and job management. Enhancing accessibility, the use of natural language in discussions guarantees that consumers may comfortably interact with the bot, irrespective of their level of computer literacy. The chatbot streamlines the user experience by combining all interactions into a single chat session. This removes the need for intricate system navigation and provides staff members with an effective and user-friendly platform to handle their job-related duties and inquiries.

2.5 QUESTION GENERATION FOR QUESTION ANSWERING

[5] In this paper, a novel neural network-based question generation method is presented, with an emphasis on utilizing large-scale question-answer pairs from Community-QA websites. The main contribution is to introduce two different approaches to question formulation. First, a retrieval-based method based on convolutional neural networks (CNN) is suggested, to retrieve questions according to the passages provided. Second, a generation-based method based on recurrent neural networks (RNN) is presented, which generates questions depending on the passages themselves.

This paper is important in ways that go beyond just creating questions. It emphasizes how these created questions might be used in real-world scenarios to improve current question-answering systems. The study presents a novel approach to improve the performance of question-answering systems by integrating the generated questions into the answer sentence selection job. The additional data acts as a signal to enhance the precision and efficacy of the procedures involved in answering questions.

The authors tested their question-generation technique extensively on three benchmark datasets: SQuAD, MS MARCO, and WikiQA to assess its effectiveness. The natural language processing community is well aware of these datasets. With the generated questions as an extra resource, the testing findings demonstrated a significant improvement in the question-answering performance. The enhancements noted in all of these benchmark datasets highlight the usefulness and possible influence of this method in improving question-answering systems for use in practical settings. Overall, through thorough experimental evaluations, this work not only introduces novel approaches to question creation but also demonstrates how they may be practically applied to improve the effectiveness and precision of question-answering systems.

2.6 LEARNING TO ASK: NEURAL QUESTION GENERATION FOR READING COMPREHENSION

[6] This study explores the field of automatically generated questions for reading comprehension tasks, with a particular emphasis on questions for sentences that are taken from text passages. A unique attention-based sequence learning model designed especially for this task is presented in the study. This study is interesting since it looks at the effects of encoding information at the sentence and paragraph levels. This model uses sequence-to-sequence learning to adopt an end-to-end trainable strategy, which deviates from previous techniques that relied on intricate natural language processing (NLP) pipelines or hand-crafted rules.

This study's main addition is the illustration of their model's efficacy in contrast to the current rule-based systems. The study team demonstrates that their method outperforms the state-of-the-art rule-based systems through extensive automatic evaluations. a development is significant because it shows how an attention-based sequence learning model may outperform traditional techniques and establish new standards for the creation of reading comprehension task questions.

To have an understanding of the qualitative elements of the created questions, the study also carries out human evaluations. Human raters determine which queries produced by their system are more naturally occurring in terms of grammar and flow. These questions are considered more difficult to answer since they deviate from the original text in both syntactic and lexical aspects. Furthermore, in comparison to questions generated by other approaches, the human assessors observe that answering these questions necessitates more reasoning. This acknowledgment of the system's capacity to generate more complex and natural-sounding questions validates its potential to raise the level of sophistication and quality of generated reading comprehension questions.

2.7 AUTOMATIC QUESTION GENERATION SYSTEM

[7] There are several complex activities involved in automating question generation, each of which has a unique set of obstacles. The main challenges in Automatic Question Generation (AQG) are identifying the target topic, choosing the question type (such as "who," "why," or "how"), and then creating the actual questions. Wikipedia frequently provides pertinent terminology or other material that is essential for inspiring queries. For each keyword category, this information can be acquired directly or by running a series of subqueries.

Nevertheless, a major challenge of current systems is the implicit character of several definition sentences taken from Wikipedia. These implied claims may make it more difficult to formulate questions accurately, which will affect the caliber and applicability of the queries that are produced. The suggested solution uses the Naïve Bayes approach in conjunction with a Supervised Learning Approach to overcome these issues. By using these methods, the system hopes to improve how well it can recognize and extract information from Wikipedia, producing definition sentences that are clearer and more instructive.

Additionally, the system incorporates Summarization and Noun Filtering algorithms, expanding its approach beyond simple definition extraction. Condensing and polishing the material that has been retrieved through summarization helps to preserve the important components required to formulate questions. Noun filtering refines the material and concentrates on key nouns pertinent to the generated queries, which helps to improve the semantic correctness of the generated questions. By combining these approaches, the information extraction process will be improved, which will result in the creation of more accurate, contextually relevant, and semantically sound questions.

2.8 QUESTION GENERATION WITH DOUBLY ADVERSARIAL NETS

[8] This paper explores the problem of generating questions on a certain topic in which labeled data is not accessible. To address this problem, the study presents DoubAN, a novel neural question-generating technique (doubly adversarial nets). The goal of DoubAN is to efficiently combine unlabeled data from the target domain with labeled data from other domains, often known as source domains. DoubAN uses two adversarial components that interact with each other during the learning process.

The question generator is assisted in learning domain-agnostic representations from the input text by the first opponent, known as the domain-classification discriminator (DC-Dis). This allows the input, irrespective of its domain origin, to be understood by the system on a broader scale. In addition to providing further training data, the second adversary—the question-answering discriminator, or QA-Dis—contributes by providing predicted reward ratings for the text-question pairs that the model generates.

Experiments carried out on datasets like SQuAD (unlabeled target domain data) and NewsQA (labeled source domain data) show that DoubAN outperforms baseline models. The results specifically show that DoubAN outperforms model variations that only use DC-Dis or QA-Dis in terms of performance. Interestingly, it is found that the relationship between DC-Dis and QA-Dis in DoubAN enhances the quality of questions generated in the target domain in a combined and indirect manner. Additionally, the study supports the efficacy and rationale of the suggested DoubAN strategy through in-depth analysis and discussion. This in-depth analysis highlights the potential of DoubAN in tackling this difficult situation and confirms the need and feasibility of using adversarial learning processes for question generation in domains without labeled data.

2.9 TEXT GENERATION FROM DATA WITH DYNAMIC PLANNING

[9] One of the fundamental challenges in language generation is data-to-text generation, or the effort of translating structured data into comprehensible, readable text. Planning the input records for text realization efficiently is one of the main challenges in this undertaking. Attempts to tackle this problem in the recent past have used static planners, which pre-plan the record structure for text production to come later. These static planners, however, are unable to modify plans in response to unforeseen or unexpected text realizations. Additionally, their adaptability and generalizability are limited by their heavy reliance on specified golden plans for supervised training.

This paper presents a novel approach that incorporates a dynamic planner to overcome these limitations. Record planning and text realization are the two separate but related steps that make up this model's breakdown of the text-generating process. The model may make revisions to plans in response to realized text thanks to the dynamic planner's iterative and flexible methodology, which also improves the model's capacity to handle unforeseen deviations during text generation. In addition, the research develops a novel likelihood-driven training approach that is customized for the planner. Instead of using annotated plans to guide the selection of input records, this technique makes use of sentence likelihood. By avoiding the requirement for explicit, hand-curated plans, this method increases the autonomy of the model and lessens its reliance on labeled data.

Apart from suggesting the likelihood-driven training technique and the dynamic planner, the study presents a new metric based on set similarity to evaluate the accuracy of the anticipated plans. This indicator helps assess how well-coordinated and effective the plans the model generates are.

The effectiveness of the suggested model is demonstrated by the experimental evaluation, which was carried out on two well-known data-to-text datasets (E2E and EPW). It shows its supremacy in producing both coherent and effective plans and high-quality text by outperforming previous works greatly on both text-based metrics and plan-based metrics. Furthermore, competitiveness in training the dynamic planner is demonstrated by the likelihood-driven training technique, confirming its efficacy and potential to improve the data-to-text production process as a whole.

2.10 SEQGAN: SEQUENCE GENERATIVE ADVERSARIAL NETS WITH POLICY GRADIENT

[10] By using a discriminative model to direct the training of a generative model, the Generative Adversarial Net (GAN) has demonstrated remarkable efficacy in producing real-valued data. Nevertheless, because the generative model's outputs are discrete, GAN becomes limited when the goal is to produce sequences of discrete tokens. This makes it more difficult to propagate gradient changes from the discriminative model to the generative model, which makes training less effective. Furthermore, only whole sequences can be evaluated by the discriminative model, which makes it more difficult to evaluate partially formed sequences and balance their current score with future predictions after the entire sequence is generated.

This research presents SeqGAN, a novel sequence-generating framework intended to get past these obstacles to handle these problems. Within reinforcement learning (RL), SeqGAN reimagines the data generator as a stochastic policy. SeqGAN uses this RL architecture to solve the generator differentiation problem, avoiding the challenge of directly conducting gradient policy updates. SeqGAN uses an RL reward signal that comes from the GAN discriminator's assessment of an entire sequence as a feedback loop for the generative model. The model is then able to optimize the sequence creation process iteratively by transmitting this signal retroactively back to the intermediate state-action phases via Monte Carlo search.

The effectiveness of SeqGAN is demonstrated using comprehensive tests on both artificial and real-world tasks. These tests show notable gains over stable baselines, showcasing SeqGAN's amazing ability to produce high-quality sequences under diverse conditions. SeqGAN shows promise as a framework to tackle the difficulties in sequence-generating tasks by utilizing both Monte Carlo search techniques and the concepts of reinforcement learning, thereby advancing the state-of-the-art in this field.

III. EXISTING SYSTEM

The technological advancements of today's fast-paced digital world have brought about a tremendous revolution in the recruitment scene. This evolution is demonstrated by the current interview chatbot system, which marks a significant transition from traditional interview approaches to an advanced automated approach. This system is a combination of advanced Natural Language Processing (NLP) algorithms and interactive conversational interfaces, specifically designed to negotiate the difficulties of applicant assessment. It was created in response to the increasing need to optimize and streamline the hiring process to meet the demands of effective candidate screening as well as the difficulties presented by a large number of applications. The current interview chatbot system is thoroughly analyzed and evaluated in this paper, which explores its strengths, weaknesses, functions, and possible areas for improvement.

- **CCViZ:** Designed to simplify the hiring process, CVViZ is a recruitment software system. It uses natural language processing to parse and filter resumes, making it easier to connect qualified candidates with open positions. The technology automates workflows like scheduling interviews and managing communications. Team coordination is improved via collaboration tools, and data-driven decision-making is made possible by analytics and reporting features. In addition to having an intuitive interface and integrating with other HR systems, CVViZ also has the potential to leverage artificial intelligence for predictive analytics and insights. Resume parsing, applicant matching, automated processes, analytics, candidate relationship management, integration possibilities, compliance, and security measures are all included in its functionality. It is recommended to visit the official CVViZ website or get in touch with the supplier directly for the most recent information.

- **Freshteam**: Designed to streamline the hiring process, Freshteam is an applicant tracking system (ATS). It automates data entry jobs with its applicant tracking, job posting, and resume parsing functions. The software offers customizable procedures, interview scheduling, and collaboration capabilities for efficient team communication. Future employment requirements can be met by creating a talent pool with the use of candidate relationship management (CRM) functionalities. For a complete hiring solution, Freshteam also provides onboarding features, an intuitive user interface, a connection with third-party applications, and tools for reporting and analytics. It is advised to check out the official Freshteam website or get in touch with the supplier directly for the most recent information.

- **Recruitment Chatbot**: By automating crucial conversations, a recruitment chatbot system simplifies the hiring process. It interacts with applicants, providing details about job opportunities and the corporate culture in addition to aiding with resume screening and natural language processing parsing. The chatbot improves the applicant experience by streamlining the application process, assisting candidates during interviews, and automating follow-ups. It handles candidate data with ease because it is integrated with applicant tracking systems (ATS). Personalized processes, data, and multi-channel assistance help streamline the hiring process, while features like compliance and feedback gathering guarantee a safe and easy-to-use experience all the way through.

IV. LIMITATIONS IN EXISTING SYSTEM

The current interview procedures face several difficulties, such as laborious manual screenings, slow response times, a lack of real-time insights, and disjointed coordination. These challenges highlight the need for a novel interview chatbot system to overcome these challenges and improve the effectiveness of candidate assessment for improved hiring procedures.

- **Massive Volume:** The number of resumes submitted for job openings on online job portals and recruitment platforms is increasing exponentially. This spike in resumes from a wide pool of individuals with a range of experiences, talents, and backgrounds is frequently linked to the convenience with which online applications may be submitted and the accessibility of these platforms.

- **Lack of Uniformity and Unstructured Information:** There is a noticeable variation in the formatting, content organization, and style of the resumes that were received. Various templates, layouts, and wording are frequently used by candidates to highlight their backgrounds and experiences. As a result, resumes don't always follow the same format or convention. making it difficult for automated systems or human recruiters to regularly retrieve pertinent data. Standardized sections of resumes are sometimes absent, which makes it challenging to locate important information—like work history, abilities, degrees, or certifications—quickly.

- **Danger of Bias and Misinterpretation**: Chatbots interpret and react to candidate inquiries using natural language processing (NLP). However, this technology can unintentionally pick up biases from the training set or misunderstand candidate questions, which could result in biased decision-making or erroneous responses. This might lead to unjust treatment or misunderstandings with applicants, which would undermine the fairness and openness of the hiring process.

- **Dependency on Training Data Quality**: The caliber and variety of a chatbot's training data have a significant impact on its efficacy. The chatbot's replies and decision-making may be less than ideal if the original dataset it was trained on is biased, lacks diversity, or is not properly representative of the applicant pool. This reliance on high-quality training data may limit the chatbot's flexibility in responding to different candidate profiles and needs, which would reduce its overall accuracy and usefulness in the hiring process.

- **Limited Complex Decision-Making:** Chatbots are great at answering simple questions and performing repetitive chores, but they may have trouble making complex decisions. For example, chatbots might not have the emotional intelligence or contextual awareness to handle delicate HR-related issues, evaluate subjective talents, or negotiate complex deals in these kinds of situations. Due to this constraint, human interaction may be required, particularly at crucial junctures in the recruiting process.

V. PROBLEM STATEMENT

Due to the large number of applications, modern recruiting processes require effective resume screening and candidate evaluation. Because it takes a lot of time to manually sort through resumes, automated recruitment systems are required. Still, there are issues with automating this procedure.

Using natural language processing (NLP) to extract structured data from a variety of resume formats is one of the main challenges. This entails selecting keywords related to programming languages from different resumes. In addition, creating and assessing questions might be difficult because multiple-choice questions are shallower than short-answer questions.

Complexity is increased when questions are created using technical ideas from several levels. Using camera monitoring to detect eye movements and stop fraud is necessary while proctoring candidates during exams.

The method attempts to effectively match employment roles with candidates' skills. Developing a strong automated recruiting system that maximizes candidate selection based on job descriptions while maintaining assessment integrity depends on finding solutions to these problems.

VI. PROPOSED SYSTEM

The suggested method is made to tackle issues like pulling programming language skills from a candidate's resume, selecting the best candidate based on a comparison of their resume and job description using various algorithms, creating questions with the help of cutting-edge natural language processing techniques, evaluating answers with BERT, and proctoring. Fig.1 displays the system block diagram. The following are the primary modules that comprise the suggested approach:

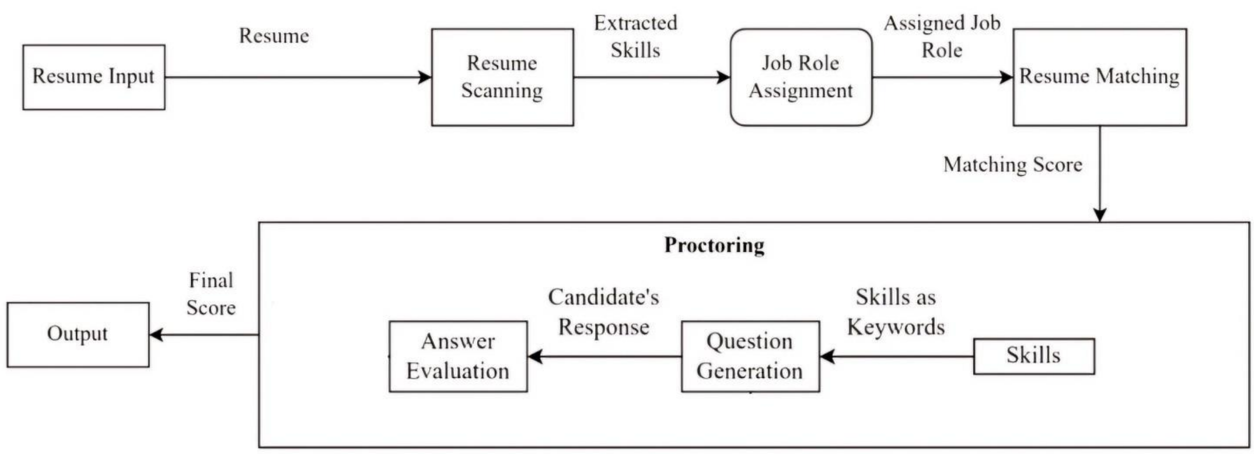

Figure 1. System Block Diagram

6.1 Resume Screening

It extracts skill parts from resumes with varying structures. The resume from file storage is accessed, and then, while parsing it, each text is stored in StringBuilder. Every word from the StringBuilder is then compared with a "skill" JSON file, which contains all the possible skills a candidate may or may not have. If any of the skill matches are found in the JSON file, then we store words in a list. Then we compare each word in the list with the skills mentioned in the job description and evaluate whether the candidate's resume is applicable for the job.

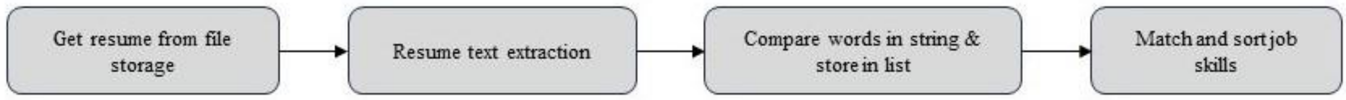

Figure 2. Resume Screening

6.2 Question Generation

In this module, the candidate will be asked questions exclusively based on the skills listed on their resume. Here we will have two sections for asking questions. Section one will have all the multiple-answer type questions and the next section will be having all the descriptive-type questions. The questions will be sorted based on the skills the candidate has on their resume and accordingly, the questions are asked.

6.3 Answer Evaluation

In this module, cosine distance and BERT embedding are used to assess the candidate's responses. Using BERT, a dense vector is created from the candidate's response that additionally considers the sentence's context. For the stored answer, an embedding of the same kind is created. After calculating the cosine similarity scores between the embeddings, the cosine distance between the two documents is found by deducting the cosine similarity of these documents from one. The results are more similar to the lower cosine distance. The average of all the similarity values is computed after this technical response evaluation module, yielding a percentage number representing the candidate's technical ability level. The database workflow where questions are fetched is depicted in Figure 3.

1. Text Similarity: This study uses text similarity to estimate how similar two documents are. In our example, the documents are the candidate's replies, which are already saved in the database. There are two types of text similarity: lexical similarity at the surface level and semantic similarity at the context level. Context level similarity takes into account both word-to-word correspondence and the document's context. Semantic similarity algorithms have successfully achieved paragraph-level similarity, which was the need for the suggested strategy. This is accomplished by creating sentence embeddings, or vector representations of documents. Next, use the cosine distance between these document features to determine any similarities between them.

2. BERT: In this case, BERT is being used to generate sentence embedding. It is primarily built on transformers that provide dense word embeddings for the input text by way of a collection of encoders. In contrast to conventional embedding approaches, this method also considers the sentence's context while embedding the sentence. It is referred to be a pre-trained bidirectional model as a result. BERT integrates semantic information into densely packed vectors. Each encoder layer generates a set of these dense vectors since BERT is quite skilled at creating them. Because BERT has been pre-trained on a sizable text corpus, it has a richer comprehension of linguistic aspects and can produce authentic embedded vectors as a final output through numerous encoder layers.

2. *ALGORITHM: Answer Evaluation*

Input: Responses submitted by the candidate Steps:

1) Generate embeddings from the documents using BERT
2) Find the cosine similarity between these embeddings
3) Calculate the similarity count for all the stored answers using the following formula

Cosine distance = 1 - Cosine Similarity where Cosine similarity = cos (x, y)

4) Calculate the mean of similarity measures calculated using cosine similarity

$$Score = \frac{1}{n}\sum_{i=i}^{n} x_i \qquad (1)$$

where n= total number of questions asked

x_i = similarity score for each answer

5) This is the final assessment score

Output: Dynamically generated and retrieved questions

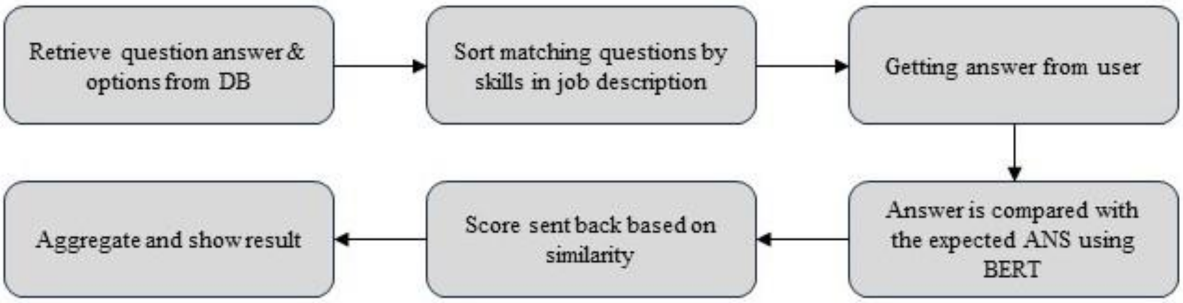

Figure 3. Database workflow

6.4 Proctoring

This module contains a DNN (Deep Neural Network), which is often used with a CNN that has undergone facial recognition training beforehand. Because online interviews and assessments are convenient, they help to reduce the malpractices that often happen. The native OpenCV libraries were imported to integrate the camera, and the VideoCapture(0) function was used to initialize the camera. In this case, 0 is supplied to access the default laptop/desktop camera. then continuously taking pictures with the camera to create a video by layering on frame-by-frame photographs. This cycle keeps going till the user completes the test. If the module detects more than one face or a cell phone, it sounds an alarm. To help identify things in the frame, the module includes a user-defined set of classes that contain a variety of objects like people, books, and cell phones. To identify and address misconduct, this module can also be useful in counting the number of faces in the visible video frame.

VII. RESULTS AND DISCUSSION

The proposed interview system is a big step towards changing the recruiting environment. This method provides an easy-to-use and accessible way to examine candidates' talents and expertise by using technology to streamline the interview process. It's critical to understand that although the system acts as a starting point for communication, it doesn't capture all of a candidate's potential.

The suggested method seeks to transform the interview process by offering a methodical but individualized way to evaluate candidates. It's crucial to remember that interviews—whether they are conducted by a human interviewer or by a cutting-edge ML system—signify the start of a candidate's career. This strategy is meant to be used in conjunction with conventional interviewing techniques, improving accessibility and speed while acknowledging the value of additional evaluations and aspects that go beyond the first conversation. Let's embrace technological innovations as catalysts for creating meaningful connections and discovering extraordinary talent as the recruitment landscape continues to change.

REFERENCES

[1] A. Zaroor, M. Maree and M. Sabha, "JRC: A Job Post and Resume Classification System for Online Recruitment," 2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI), 2017, pp.780 787, doi: 10.1109/ICTAI.2017.00123.

[2] Pradeep Kumar Roy, Sarabjeet Singh Chowdhary, Rocky Bhatia, A Machine Learning Approach for automation of Resume Recommendation system, Procedia Computer Science, Volume 167, 2020, Pages 2318-2327, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2020.03.284.

[3] International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 06 Issue: 04 | Apr 2019 "Interactive Interview Chatbot" Shriniket Yakkundi[1], Amey Vanjare[2], Vinay Wavhal[3], Shreya Patankar[4].

[4] S. Tadvi, S. Rangari, and A. Rohe, "HR Based Interactive Chatbot (PowerBot)," 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA), 2020, pp. 1-6, doi: 10.1109/ICC-SEA49143.2020.9132917.

[5] Nan Duan, Duyu Tang, Peng Chen, and Ming Zhou. 2017. Question Generation for Question Answering. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, pages 866–874, Copenhagen, Denmark. Association for Computational Linguistics.

[6] Su, MH, Wu, CH, Huang, KY, Hong, QB Huang, HH 2018, Learning to Ask: Neural Question Generation for Reading Comprehension Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH, vol. 2018-September, pp. 1006-1010. https://doi.org/10.21437/Interspeech.2018-1007.

[7] P. Pabitha, M. Mohana, S. Suganthi and B. Sivanandhini, "Automatic Question Generation system," 2014 International Conference on Recent Trends in Information Technology, 2014, pp. 1-5, doi: 10.1109/ICR-TIT.2014.6996216.

[8] J. Bao, Y. Gong, N. Duan, M. Zhou, and T. Zhao, "Question Generation With Doubly Adversarial Nets," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 26, no. 11, pp. 2230-2239, Nov. 2018, doi: 10.1109/TASLP.2018.2859777.

[9] S. Yang, Y. Liu, D. Feng, and D. Li, "Text Generation From Data With Dynamic Planning," in IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 30, pp. 26-34, 2022, doi: 10.1109/TASLP.2021.3129346.

[10] Lantao Yu, Weinan Zhang, Jun Wang, and Yong Yu. 2017. SeqGAN: sequence generative adversarial nets with policy gradient. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI'17). AAAI Press, 2852–2858.

[11] UGC Approval 14 June W.e.f of CARE List UGC Approved Journal no 63975 Preeti Bhat, Harsha Pariyani, Anshika Sinha, Roshni RoteProf. N. A. Mulla.