Enhancing Recruitment Processes Through Integration of Personality Traits and Professional Skills Analysis

Maleesha De Silva
Department of Information Technology
Specializes in Data Science
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
maleeshadesilva21@gmail.com

Sandani Zoysa
Department of Information Technology
Specializes in Data Science
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
sandani991012@gmail.com

Manushika Maldeniya
Department of Information Technology
Specializes in Data Science
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
manushikamaldeniya1999@gmail.com

Shanali de Silva
Department of Information Technology
Specializes in Data Science
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
shanalids6@gmail.com

Dr. Anuradha Karunasena
Department of Information Technology
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
anuradha.k@sliit.lk

Dr. Lakmini Abeywardhana
Department of Information Technology
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
lakmini.d@sliit.lk

Abstract— **The success and growth of an organization are heavily dependent on the selection of the right candidates. The conventional approach of relying solely on resumes and academic backgrounds falls short in adequately evaluating a candidate's suitability for a job role. This research paper proposes a novel method of candidate recruitment that streamlines the process by integrating advanced machine learning algorithms and natural language processing techniques. This study introduces a new approach to measuring programming language proficiency by analyzing GitHub data and comparing candidates using a knowledge graph. LinkedIn skill data is used to predict job categories through a SVM multiclass classification machine learning model. Additionally, Referee feedback is evaluated using sentiment analysis to improve recommendations. The study also utilizes K-means clustering and NLP techniques to identify the Big Five personality traits distribution among candidates and assess their alignment with specific job roles. The XGBoost model shows the best performance in predicting the personality cluster of candidates. The findings of this study contribute to the growing field of HR technology and provide valuable insights for optimizing recruitment practices in the IT industry using machine learning.**

Keywords—Big Five traits, Machine Learning, Natural Language Processing, Professional Skills, Recruitment, Reference checking, Sentiment analysis

I. Introduction

Recruiting the most suitable candidates is a crucial process for any organization striving for success. However, when selecting candidates for IT job positions, decisions cannot solely rely on their resume and academic background. It's essential to consider their experience in relevant areas, proficiency in programming languages, and how well their personality aligns with the job role.

Understanding and identifying personality attributes of employees in the workplace is essential for creating a positive and productive work environment. According to British psychologist Adrian Furnham[1], personality traits play an important role in an individual's success at work, and studies show that personality is more predictive of professional effectiveness than previous experience.

Professional skills, also called technical skills, are a type of skill we use in the workplace. These skills cover all the knowledge and experience we specifically use to do our jobs[2]. Hiring candidates with the right professional skills is a good way to ensure high job performance.

Although many recruiters are familiar with modern technology tools and their benefits, not all of them effectively utilize technology to analyze candidate data and pinpoint the most fitting individuals for specific roles [3]. Nevertheless, analyzing data from diverse sources like GitHub and LinkedIn, referee feedback on candidates along with the assessment of personality traits, can significantly enhance the identification process [4]. The drawback, however, is that this approach can be time-consuming, leading some companies to abandon the use of such data.

To address these challenges, this research paper proposes a machine learning-based approach that streamlines the recruitment process by leveraging professional skills and personality traits to identify the most suitable candidates for specific job roles.

The 'Big Five Model' was chosen as the personality evaluation model in this research due to its strong relevance to workplace performance. Research from the Academy of Management[5] indicates that The Big Five personality traits can influence job performance outcomes. Different job roles and tasks can trigger specific traits, and knowing which individuals possess compatible traits for a position allows you to create a high-performing team.

Candidate's professional skills are predicted by scraping and analyzing data from platforms such as GitHub, LinkedIn, and reference feedback forms. This process aims to provide recruiters with a comprehensive understanding of the candidates' abilities and talents prior to the recruitment.

The study's findings carry significant implications for HR technology, as integrating machine learning and natural language processing techniques can lead to improved candidate selection, reduced time-to-hire, and overall enhanced recruitment outcomes within the IT industry.

II. Literature Review

In recent years, numerous studies have focused on exploring methods to go beyond evaluating only CVs and academic skills during candidate recruitment, with the aim of optimizing the overall recruitment process.

The research paper titled "DevFlair: A Framework for Streamlining the Pre-screening Process in Software Engineering Job Applications" authored by Jayasekara et al.[6] presents a novel system designed to automate the evaluation of job candidates' suitability. The system leverages data from various sources such as social media,

979-8-3503-5813-1/23/$31.00 ©2023 IEEE

GitHub, and open-ended questionnaires to predict the compatibility of candidates for specific job positions. In this study, the authors use a Personality Prediction Model (PPM) to forecast the Big Five personality traits of candidates. The PPM utilizes data extracted from the candidates' LinkedIn profiles through the ProxyCurl LinkedIn API. This data is processed and fed into a predictor to receive a compound probability score, which serves as the final output representing the candidate's personality.

According to Gajanayake et al.[7], they propose a screening process that utilizes phone call transcripts containing candidate responses to open-ended questions to assess their personality traits. The study uses feature extraction techniques to identify the Big Five traits of candidates from the responses. Multiple supervised learning classification algorithms are utilized to evaluate the model, and it is found that Logistic regression achieves the highest level of accuracy. Also, this research has proposed a pre-screening solution to screen the applicants for the position of Software Engineer where candidates are screen-based GitHub profile data along with CV data, matching the total number of LinkedIn profile skills with job position with academic transcripts and, Recommendation letters. For recommendation letters they have applied a multi-class classification machine learning approach to a candidate as recommend, highly recommend, neutral, and weakly recommend.

Personality traits can be described as patterns of thought, feeling, and behaviors that are characteristic of an individual. Aligning candidate personality traits with job role requirements is important as it ensures employees possess the necessary skills for success, promotes better job fit, and boosts job satisfaction and performance. Additionally, it reduces employee turnover benefitting both the organization and the employee. Several studies have examined the relationship between personality traits and job performance consistently indicating a strong correlation between the two.

Studies have found that certain personality traits are more important for some jobs than others. Stephen P. and Timothy A.[8] suggest that conscientiousness has the most significant influence on job performance. Employees with high conscientiousness tend to be knowledgeable, organized, and reliable. They are also more likely to be leaders and take on additional responsibilities. Employees who are high in neuroticism are more likely to experience stress, anxiety, and burnout. They may also be more likely to take sick leave. Extroverts are more likely to be outgoing, assertive, and enthusiastic. They are also more likely to be willing to take on leadership roles and be involved in social activities. Employees who are open to new experiences are more likely to be creative, innovative, and adaptable. They are also more likely to be supportive of change. Agreeable employees are more likely to be cooperative, helpful, and friendly. They are also more likely to be satisfied with their jobs.

According to M. Aqeel Iqbal et al.[9] researchers have identified that the Big Five Model (BFM) also known as the Five-factor Model is the most prominent and trusted personality classification model.

Dr. S.K. Nivetha et al.[10] conducted a study that uses the facial-width-to-height ratio (fWHR) to predict personality traits in interview candidates. The researchers utilize a Convolutional Neural Network (CNN) to extract the fWHR from candidate images. Subsequently, this fWHR measurement is used to predict a person's personality traits with an impressive accuracy rate of 92%.

The paper [11] explores the use of GitHub REST API to gather information on potential candidates from their user profiles. This research paper discusses how using novel methods to inspect public GitHub repositories can help manage the influx of candidate profiles during the recruitment process.

A research conducted in 2022 [6] to automate the prescreening process of job candidates used TSUAM (Technical Skill Usage Analysis Model) which focuses on creating a scoring system to assess technical skills usage on GitHub based on the number of repositories.

Another interesting research was been done in 2021 [12] to predict user interest based on social network profile data using machine learning. This research paper outlines a system designed to analyze data from Facebook and Twitter by utilizing scraping tools to assess user interests, offering public opinion and gaining insights into people’s sentiments.

When it comes to Employee reference checking, [13], [14] mentions reference checking in the hiring process can be hugely helpful for choosing the most suitable candidates by confirming that the candidates have all the professional skills required or uncovering something disappointing.

The current research on employee personality traits focuses primarily on identifying individual traits rather than assessing their suitability for specific job roles. This research aims to bridge this gap by considering the compatibility between an employee's traits and a particular job role. In recent years’ competitive job market, identifying the right candidate for a job is crucial for an organization's success.

While traditional hiring processes rely heavily on resumes and interviews, these methods may not be enough to fully evaluate a candidate's professional skills. In this research, we aim to critically analyze the existing literature on the use of LinkedIn, GitHub profiles, and Reference check forms in evaluating professional skills during the candidate hiring process.

In the following sections, we will discuss the methodology followed in developing the proposed solution, the data collection and preprocessing procedures, the machine learning algorithms utilized, and the evaluation metrics utilized to assess the performance. We will then present and discuss the study's results, followed by the conclusion and suggestions for future research.

III. METHODOLOGY

A. Aligning Candidate Personality Traits with Job Role Requirements

The personality assessment model used in this component to assess the candidate-job fit is the ‘Big Five Personality Model’.

To assess the alignment between a candidate's personality and a specific job role, this study employs a hybrid approach, combining the results from a multiple-choice questionnaire (MCQ) and an open-ended questionnaire. These collected results are subsequently integrated with the job's desired personality traits to determine the level of compatibility between the candidate and the job role (Figure 1).

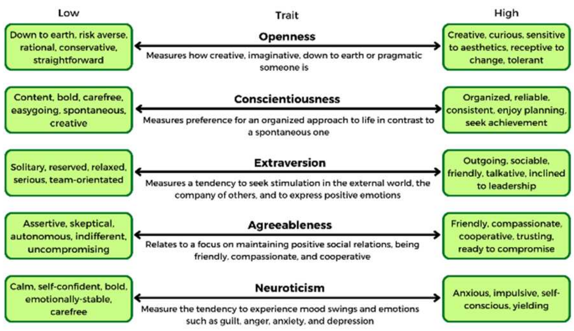

Fig. 1. The Big Five Personality Traits

This component utilizes two datasets. The first dataset comprises candidate responses to 50 self-rating questions, which were sourced from Kaggle[15]. Each question addresses a specific personality trait, with a total of 10 questions per trait. Participants provided ratings for their responses on a scale of 1 to 5. The other dataset was prepared manually by collecting candidate responses to five open-ended questions corresponding to the Big Five Traits.

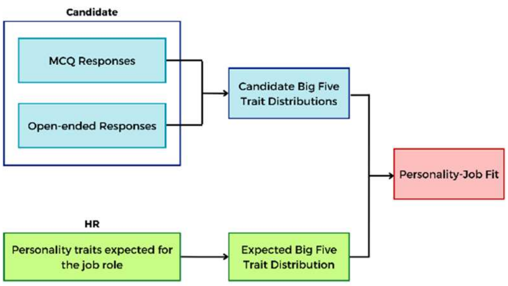

Fig. 3. Overview diagram - Candidate Personality Analysis

1) *MCQ Questionnaire*

The responses in the MCQ dataset are used to identify potential personality clusters that a candidate may belong to. To achieve this objective, the K-means Clustering Algorithm was chosen. The optimal number of clusters (k value) is determined using the Elbow visualization technique. Given that the number of clusters is known in advance, and considering the scale of the dataset, the K-means Clustering algorithm emerges as the most suitable approach for this task [16]. Using the clustering algorithm, the cluster to which each response in the dataset belongs to is predicted. These predictions along with the responses will be used to create a supervised dataset for future steps, enabling us to validate the predictions. The candidate responses for the MCQ questionnaire were processed to derive a distribution of the Big Five Personality traits for each candidate.

2) *Open-ended Questionnaire*

As the initial step, keyword extraction was performed using three different methods: the Bag of Words Model, the TF-IDF Vectorizer Model, and the keyBERT Model. Among these methods, the TF-IDF model demonstrated the ability to extract the most relevant keywords for each trait from the candidate responses. 40 keywords were extracted for each trait. Table I below shows a sample list of keywords compiled for each trait based on the extracted keywords.

TABLE I. KEYWORD LIST

Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism
new	work	new	understand	problem
risk	time	people	feedback	stress
life	task	social	improve	difficult
learn	complete	relationship	listen	manage
believe	accurately	connection	ask	break
....	....	....	....	....

This keyword list serves as the benchmark against which candidate responses are compared. The candidate responses for the open-ended questionnaire will undergo preprocessing following the aforementioned steps, and keywords will be extracted. These extracted keywords will be compared against the previously prepared keyword list (see Figure I) to generate a distribution of the Big Five traits for the candidate. By combining the results from the multiple-choice questionnaire and open-ended questionnaire, we can obtain a comprehensive Big Five trait distribution for each candidate.

3) *Determining Candidate-Personality Job Fit*

The candidate's desired personality traits for a particular job role are considered. Figure 3 below shows how various personality traits are mapped to the Big Five traits[17] to determine the expected distribution of Big Five traits for that role.

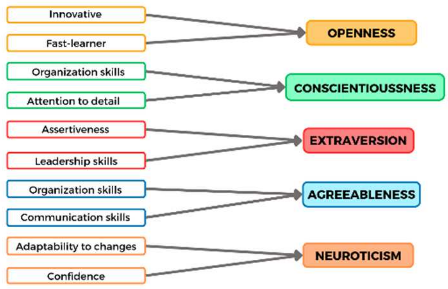

Fig. 3. Mapping personality traits with the Big Five Traits

The candidate's distribution of the Big Five traits is assessed and compared to the job role's personality requirements. This evaluation produces a score indicating how closely the candidate's personality aligns with the job role.

B. *Evaluating Professional Skills using Digital Footprints and Reference check forms.*

1) *GitHub Code Proficiency Comparator*

GitHub user profile data was used to evaluate Programming language proficiency in this research.

To evaluate programming language proficiency, GitHub user profile data was utilized in this research. The GitHub REST API, along with the PyGitHub modules, facilitated data retrieval from the candidates' profiles by leveraging a personal access token for authentication through the API's GET request endpoint, **'/users/{username}/repos'**.

The user's repositories associated with each candidate were fetched. Iterating through the repositories in the API response, a repository object was generated for each repository using a loop. The **'get_languages()'** method provided by the PyGitHub modules enabled the retrieval of repository language proficiency, represented by a dictionary where the keys denoted the languages used and the values indicated the number of bytes written in each language. By accumulating line counts for each language across all repositories, the overall language proficiency was calculated. This proficiency was determined based on the number of lines of code written in each language across all repositories, ultimately displayed as a percentage representing the candidate's overall programming language proficiency.

This approach considers the repository count as the weight, assuming that a user with more repositories may have a broader range of language proficiency. The methodology allowed for the inclusion of multiple candidates, facilitating a comparison of their programming language proficiencies.

2) *LinkedIn Skill-Based Job Title Classifier*

LinkedIn profile data was extracted using a third-party paid API service called ‘ProxyCurl' to retrieve candidate skills added to the LinkedIn profile under "skills". The candidate's LinkedIn profile must be public to extract the skills for prediction.

A machine learning approach was used with a dataset that includes job roles and categories specific to the IT industry and the required skill set for each. This dataset was created including IT industry-related job roles, categories and required skills by conducting a survey. A multiclass classification approach was used to predict the matching or classifying job candidates into specific job categories.

In order to perform text classification and predict the relevant job category based on skills mentioned in LinkedIn profiles, several steps were undertaken. Firstly, the text data was processed by transforming it into vectors using the Term Frequency-Inverse Document Frequency (TFIDF) method, employing functions such as min_df, sublinear_tf, and stop_words to enhance the quality of the vectorization. Following this, various multi-classification models, including RandomForest, LinearSVC, MultinomialNB, LogisticRegression, and XGBClassifier, were explored. Cross-validation techniques and hyperparameter tuning were applied to assess the accuracy of each model, ultimately selecting the most effective supervised machine learning model. Finally, the text classification model was evaluated using the extracted skills from LinkedIn scraping, enabling the prediction of job titles related to the identified skills.

C) Sentiment Profiling for Employee References

In this research endeavor, an analysis was conducted to ensure the candidate has the professional skills needed for the role the employer is trying to fill. For this, a questionnaire approach from referees was used. Two questionnaires were set up as a university reference check form and an employee reference check form to discover professional skills of a candidate.

A reference checking form was sent to the relevant parties as mentioned in the resume. As a time-saving approach in a flood of candidates, sentiment analysis was applied without going through each.

Feedback was extracted and applied necessary text processing steps utilizing NLTK libraries. Then performed a sentiment analysis using **'SentimentIntensityAnalyzer'** from NLTK, using polarity scores. Then sentiment scores were used to visualize the sentiment using a pie chart created with Matplotlib. Furthermore a ‘word cloud’ was created to visualize the most frequent words in the extracted feedback as it helps in quickly understanding the key theme and most prominent words in the feedback.

IV. Results and Discussion

A. Aligning Candidate Personality Traits with Job Role Requirements

1) *MCQ Questionnaire*

The Elbow Visualization technique revealed a k-value of 5 indicating that the responses within the dataset can be grouped into 5 distinct personality clusters.

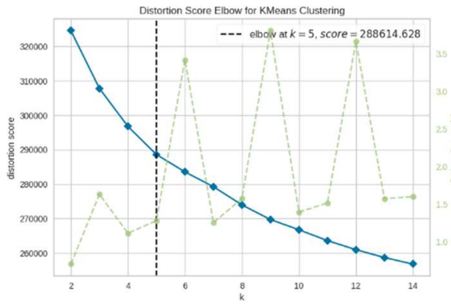

Fig. 4. Elbow Visualization to determine the optimal k-value for k-means clustering

The Big Five trait distribution for the clusters is shown below.

TABLE II. BIG FIVE TRAIT DISTRIBUTION FOR THE 5 PERSONALITY CLUSTERS

	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism
Cluster 0	3.5	3.3	3.1	3.3	2.9
Cluster 1	3.0	2.9	2.9	3.0	3.2
Cluster 2	3.4	3.35	3.1	3.33	3.9
Cluster 3	0.35	0.38	0.43	0.40	0.37
Cluster 4	3.2	2.9	3.0	3.0	2.2

Identifying which cluster, a candidate belongs to would help recruiters to determine what kind of personality a candidate has. All of the responses in the dataset are assigned a cluster using the algorithm to create a supervised learning dataset.

The supervised learning dataset obtained from the MCQ response dataset is subjected to validation using RandomForest, Naïve-Bayes, and XGBoost algorithms. The performance parameters of the algorithms are shown below.

TABLE III. MODEL PERFORMANCE PARAMETERS - PERSONALITY PREDICTION

Model	Accuracy	Precision	Recall	F1 Score
RandomForest	0.99	0.992	0.992	0.992
Naïve-Bayes	0.96	0.958	0.97	0.97
XGBoost	0.95	0.96	0.96	0.96

Based on the results obtained above, the XGBoost algorithm is the most suitable to predict the personality cluster that a candidate belongs to.

2) *Open-ended Questionnaire*

The keywords from the candidate responses will be compared against the prepared keyword list shown in Table I to obtain a Big Five trait distribution for each candidate (Table IV).

TABLE IV: BIG FIVE TRAIT DISTRIBUTION FOR A CANDIDATE.

Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism
3.1	2.9	2.3	3.35	2.2

3) *Determining the Candidate-Personality Job Fit*

A radar graph is utilized to visually represent and compare the candidate's Big Five trait distribution with the expected Big Five trait distribution for the job role. This visualization provides recruiters with a quick and intuitive understanding of how well the candidate aligns with the requirements of the job role.

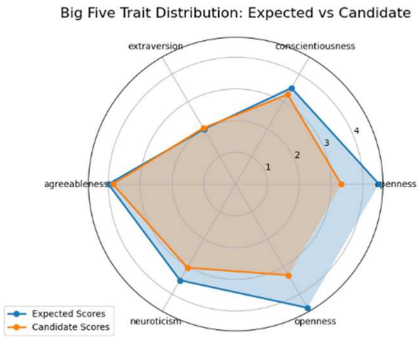

Fig. 5. Expected vs Candidate Personality Comparison

The compatibility between the candidate’s personality and the expected personality traits is calculated by the equation shown below:

$$\text{Score} = \Sigma\left(\frac{trait(candidate)}{trait(expected)}\right) * \frac{100}{5}$$

The sum of ratios between the candidate and the expected scores for each of the Big Five traits is taken as a percentage. As neuroticism is negatively correlated with job performance[6], the probability that the candidate does not possess the neuroticism personality trait is considered.

The most closely related research conducted previously [6] employed 5 distinct machine learning models to identify the Big Five traits. However, the accuracy achieved by each of these models was lower compared to the accuracy attained by the XGBoost model discussed in this paper, thus demonstrating significant improvement over the previous research. Moreover, the capability to evaluate the alignment between the candidate and their personality represents a notable advancement beyond the scope of prior related studies.

B. *Evaluating Professional Skills using Digital Footprints and Reference Letters.*

1) *GitHub Code Proficiency Comparator*

This section explores the extraction and analysis of data from GitHub profiles to calculate the weighted language proficiency scores by multiplying the language proficiency value by the reciprocal of the user's repository count. By evaluating a single candidate, the result presents a comprehensive summary of the candidate's programming language proficiencies, represented as percentages as in Figure 6. The approach provides a quantitative measure of the candidate's overall programming language skills, leveraging data from their GitHub profile.

```
Overall Language Proficiency (Percentage):
- Java: 76.55%
- JavaScript: 5.38%
- HTML: 10.61%
- CSS: 7.45%
```

Fig. 6. Single candidate's overall language proficiency

In the context of comparing multiple candidates, the aforementioned results for each candidate are visually depicted using a knowledge graph as in Figure 7, nodes represent users and edges represent the language proficiency, with the weight attribute indicating percentage contribution. for common languages for each candidate, as illustrated in the accompanying image. This knowledge graph serves as a graphical representation that enables logical comparisons between candidates based on their programming language proficiencies. By leveraging this knowledge graph, one can systematically evaluate and contrast the language proficiencies of multiple candidates, facilitating informed decision-making in the candidate selection process.

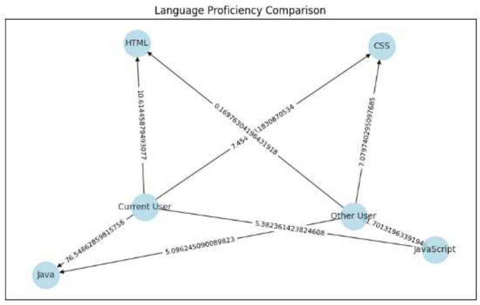

Fig. 7. Knowledge graph comparison between peer candidates

2) *LinkedIn Skill-Based Job Title Classifier*

For predicting the job category, the multiclass classification approach was used. Table V depicts the Linear Support Vector Machine outperforming all the other classification algorithms giving an accuracy of 66% as shown in the below table.

TABLE V: ACCURACY SCORES FOR EACH MODEL

Model	LinearSVC	Logistic Regression	MultinomialNB	RandomForest Classifier	XGBClassifier
Mean Accuracy	0.664	0.579	0.550	0.579	0.437
Standard Deviation	0.131	0.143	0.109	0.119	0.096

The LinkedIn REST API resulted in scraping the candidate's skills mentioned in the LinkedIn profile. Combining the above two results one can predict the most suitable skills-based job position. Predicting job categories based on skills brings efficiency, personalization, and accuracy to the hiring process, benefiting both employers and candidates by aligning skills with job roles and optimizing resource allocation.

3) *Sentiment Profiling for Employee References*

The findings were thoughtfully visualized to ensure clarity and comprehension. A visually appealing pie chart representing the sentiment analysis results, with segments for positive, negative, and neutral sentiments, along with their respective percentages as in Figure 8. Word cloud, generated using the WordCloud library, artfully presented as in Figure 8 the most frequently occurring words in the analyzed text, thereby accentuating significant themes and keywords associated allowing a succinct representation.

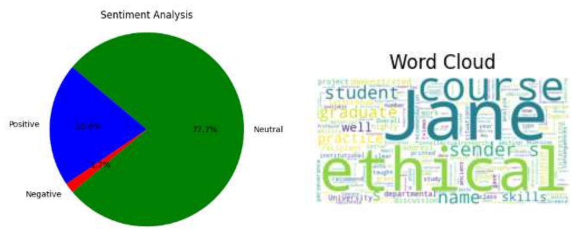

Fig. 8. Sentiment results for reference feedback

Further validation of the sentiment analysis can be achieved through a manual assessment facilitated by human intervention. This approach involves enlisting human annotators to evaluate and assign sentiment labels to a subset of the analyzed text. By comparing the human-assigned labels with the sentiment scores generated by the sentiment analysis tool, a comprehensive assessment of the tool's accuracy and efficacy can be ascertained.

Comparing with Gajanayake et al.[6], machine learning categorical prediction on recommendation letters, it is more likely most recommendation letters were given with positive feedback, This novel method with questionnaire covers strengths and weaknesses of the candidate which helps to gain a sentiment expressed in feedback that carries responses with well-informed analysis in decision-making.

V. Conclusion and Future Work

This research aims to optimize the job recruitment process in the IT industry by shifting towards automated approaches. It focuses on the impact of automation in hiring freshers and recent graduates, tailoring the practices to this segment of candidates. The proposed approach combines personality trait assessment, GitHub, and LinkedIn profile analysis to provide a comprehensive view of candidates' skills and traits. This innovative method contributes to more accurate hiring decisions. In future work, the inclusion of techniques to validate the information provided in candidate LinkedIn profiles and Questionnaire responses will significantly improve the accuracy of the proposed solution. This enhancement will prove advantageous for organizations seeking to identify the most fitting candidates for their workforce.

VI. References

[1] "Which Personality Attributes are Most Important in the Workplace," May 19, 2023. https://www.thomas.co/resources/type/hr-blog/which-personality-attributes-are-most-important-workplace

[2] J. Gain, "Why Valuing Professional Skills is so Important," Feb. 25, 2021. https://management30.com/blog/valuing-professional-skills-important/

[3] S. Karpe, "The Top Recruitment Challenges of IT Services Companies," Jun. 02, 2020. https://blog.imocha.io/the-top-recruitment-challenges-of-it-services-companies

[4] S. Tennakoon, R. Islam, D. Wijerathna, R. Deen, P. Sumathipala and L. Abeywardhana, "An Interactive Application for University Students to Reduce the Industry-Academia Skill Gap in the Software Engineering Field," 2023 4th International Conference for Emerging Technology (INCET), Belgaum, India, 2023, pp. 1-6, doi:10.1109/INCET57972.2023.10170591.

[5] T. A. Judge and C. P. Zapata, "The Person–Situation Debate Revisited: Effect of Situation Strength and Trait Activation on the Validity of the Big Five Personality Traits in Predicting Job Performance," *Acad. Manage. J.*, vol. 58, no. 4, pp. 1149–1179, Aug. 2015, doi: 10.5465/amj.2010.0837.

[6] R. T. R. Jayasekara, K. A. N. D. Kudarachchi, K. G. S. S. K. Kariyawasam, D. Rajapaksha, S. L. Jayasinghe, and S. Thelijjagoda, "DevFlair: A Framework to Automate the Pre-screening Process of Software Engineering Job Candidates," in *2022 4th International Conference on Advancements in Computing (ICAC)*, Colombo, Sri Lanka: IEEE, Dec. 2022, pp. 288–293. doi: 10.1109/ICAC57685.2022.10025337.

[7] R. G. U. S. Gajanayake, M. H. M. Hiras, P. I. N. Gunathunga, E. G. Janith Supun, A. Karunasenna, and P. Bandara, "Candidate Selection for the Interview using GitHub Profile and User Analysis for the Position of Software Engineer," in *2020 2nd International Conference on Advancements in Computing (ICAC)*, Malabe, Sri Lanka: IEEE, Dec. 2020, pp. 168–173. doi: 10.1109/ICAC51239.2020.9357279.

[8] S. P. Robbins and T. A. Judge, *Essentials of Organizational Behavior*, 14th ed. Pearson, 2017.

[9] M. A. Iqbal, F. A. Ammar, A. R. Aldaihani, T. K. U. Khan, and A. Shah, "Building Most Effective Requirements Engineering Teams by Evaluating Their Personality Traits Using Big-Five Assessment Model," in *2019 IEEE 6th International Conference on Engineering Technologies and Applied Sciences (ICETAS)*, Kuala Lumpur, Malaysia: IEEE, Dec. 2019, pp. 1–5. doi: 10.1109/ICETAS48360.2019.9117561.

[10] S. K. Nivetha, M. Geetha, R. S. Latha, K. Sneha, S. Sobika, and C. Yamuna, "Personality Prediction for Online Interview," in *2022 International Conference on Computer Communication and Informatics (ICCCI)*, Coimbatore, India: IEEE, Jan. 2022, pp. 1–4. doi: 10.1109/ICCCI54379.2022.9740980.

[11] S. Gupta, B. Gupta, and S. Gupta, "A Novel Method for Technical Candidate Assessment using Github Repository Inspection Automation," in *2022 International Conference on Computer Communication and Informatics (ICCCI)*, Coimbatore, India: IEEE, Jan. 2022, pp. 1–5. doi: 10.1109/ICCCI54379.2022.9740986.

[12] H. K. S M, S. Hegde, S. G, S. M, S. R, and S. L. N, "User Interest Prediction based on Social Network Profile with Machine Learning," in *2021 6th International Conference for Convergence in Technology (I2CT)*, Maharashtra, India: IEEE, Apr. 2021, pp. 1–6. doi: 10.1109/I2CT51068.2021.9418126.

[13] "How Do You Conduct A Great Reference Check?" https://www.personio.com/hr-lexicon/applicant-screening/

[14] B. M. Anderson, "13 Questions You Should Ask When Checking References," Jun. 23, 2021. https://www.linkedin.com/business/talent/blog/talent-acquisition/reference-check-questions

[15] B. Tunguz, "Big Five Personality Test." 2020. [Online]. Available: https://www.kaggle.com/datasets/tunguz/big-five-personality-test

[16] C. Ellis, "When to use K-means clustering." https://crunchingthedata.com/when-to-use-k-means-clustering

[17] J. Darby, "HR Blog," *thomas.co*, May 16, 2023. https://www.thomas.co/resource-type/hr-blog