

Improving Resume Screening with NLP and Machine Learning: Addressing Efficiency and Fairness

Saisree Pulavarthi¹, B. Rahul Reddy², Thatipamula Sairam³, Abhishek Bhattacharjee⁴ and Sathyasaketh jallipalli⁵

¹⁻⁵Lovely Professional University Punjab, India

Email: saisree.pulavarthi@gmail.com, boggetirahulreddy@gmail.com, sairamthatipamula69@gmail.com, abhishek.27306@lpu.co.in, jallepallisathyasaketh@gmail.com

Abstract—The demand for logistical tools such as automated resume screening has grown in response to the need of organizations to streamline their hiring funnel. This approach of using manual review process is labour-intensive and eschews a potential for bias that can solned to both inefficiency and unfairness in their hiring decisions.

The focus of this paper lies in using natural language processing (NLP) along with Machine Learning(ML) to improve the pace and fairness of different resume screening systems. We present a scalable and automated solution to make the hiring process more fair by reducing bias in resumes using NLP techniques or extracting useful information from each resume & develop ml models to analyse profile of candidates.

The research addresses critical problems which are to process the unformalized data coming in from resumes, to deal with algorithmic biases and also model performance for a broad range of jobs.

We also delve into the questions of data and ethics in how AI can be used for recruitment, as well discuss approaches to help are bias from surfacing. These numbers speak for themselves, it demonstrates that by using NLP with ML on both ends makes recruiting a breeze and in addition driving increased fairness at the same time finally intended again: job-related credentials can be de-emerged (candidate focus) without losing proper measures of quality. Those findings provided an aerial view of what should come next for ensuring AI-powered talent and recruitment systems operate more blandly. Submit a job and in the contemporary job market is generating hundreds of applications that must be reviewed fairly.

Index Terms— NLP, screening, Resume, Ai, Application, ml, ATS, smart resume, resume filter.

I. INTRODUCTION

In recent times, Manual resume screening the traditional way is labour-intensive and it leaves room for bias leading to unfair hiring practices. With the booming demand for quicker and unbiased hiring processes, organizations are turning towards automated solutions to improve decision-making and optimize recruitment workflows.

With automation, technological trends such as Natural Language Processing (NLP) and Machine Learning(ML), have proved to be extremely effective in speeding up the process of resume screening. NLP can extract useful information from unstructured text on resumes, while the ML models could evaluate candidates and rank them based on qualification scores to avoid any bias of humans adding or reducing the probability that an applicant will be hired.

Nevertheless, these technologies can be difficult to adopt because we need to handle variation in the formats of resumes; reduce algorithmic biases; and maintain a superior performance for various roles across different industries. In this paper, we analyse how to combine NLP and ML related technologies in order to build a more efficient, scalable and less biased system for resume screening. The solution is designed to more efficiently (and equitably) recruit talent by eliminating the barriers inherent in our current process and enabling an uncluttered assessment of a candidate's qualifications based on their experiences. Additionally, the study then examines ethical considerations of AI recruitment and methods to mitigate bias while being transparent. This study informs the future design of equitable and nimble recruitment systems that accommodate change over time in the labour market.

II. LITERATURE REVIEW

The project seeks to improve recruitment screening in a telecommunications company with the influence of historical data through predictive analytics. It uses machine learning algorithms, such as Random Forest (RF), preferably fused with Naïve Bayes classifier to predict candidate acceptance and decrease recruitment expense/time per hire via enhanced candidate quality. So, historical data of the previous 4 years is collected and cleaned to train models on labelled data showing accepted/not-accepted containing details about an international student. Model Validation: Using cross-validation technique on Metrics like Accuracy, Precision & Recall. The analysis shows that the Random Forest model is doing a better job, with an accuracy of 71.5%, and Recall 83.4%. [1].In the study that follows, we contextualize modern employee recruitment strategies and offer a coherent perspective on how they are being impacted by technology and emerging digital trends. It examines what these changes mean for job seekers and companies, Results reveal that the e-recruitment very well increases efficiency and reduces cost to a great extent along with giving an impression of future growth potential (21st-century image) over tech-savvy job seekers. Some of the challenges are that an article is not it's indication and PSUCs can be scratchy, blah etc. [2].

Ketla, Daniel J. “Research on the Variegated Impact of Artificial Intelligence (AI) in Human Resource Management: A Study with Reference to Hire and Fire Operations as well as Motivation.” The illustration of AI to really increase efficiency, reduce bias and inform strategic decision making. Core elements comprise an assessment of AI-based recruitment solutions, identification and elaboration on prospects as well hurdles like ethical considerations or protection of personal data, case studies to offer practical advices from fruitful deployments of AI. [3]. Using blockchain in the study is expected to improve data authentication, simplify recruitment process and help with validated qualification-driven enhanced employee placement. Basic methods include a blockchain voting platform for candidates, automatic change verification systems and consensus-sustained data incorruptibility to inhibit unsanctioned changes. Data collected to date also shows an improvement in data resiliency, greater efficiency and improved trust/transparency within the recruitment process as a result of easing skill match which ESCO helps enormously with particularly for smaller companies. [4].

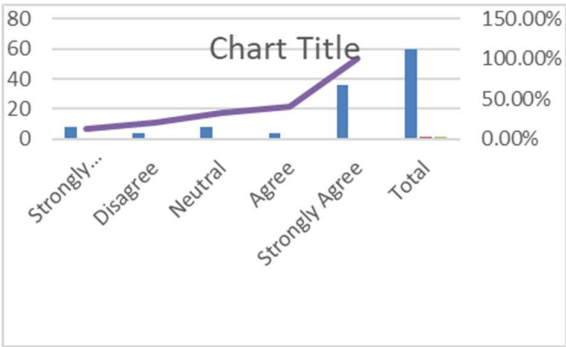


Figure 1. Represents the comparison of above Table. II

In this study, we analytically analyse how digital technologies such as artificial intelligence (AI) can leverage global talent identification, selection and retention by strategically matching organizational values with those of candidates. It offers a smart recruiting solution that examines candidate profiles from numerous online resources, provided and supports diverse assessments to prevent gender-based bias in terms of ages or nationalities. These tools range from a multiple case study framework, Grounded Theory to Participant Observation and Literature

Review or Content Analysis as well as Big Data Analytics since Talent movements can be visualized by the hundreds per candidate. [5]. In this study, we conceptually explore the way in which digital technologies like artificial intelligence (AI) will enable global talent identification, selection, and retention to be altered by connecting organizational values with those of potential candidates. An AI powered smart recruitment solution that reviews candidate profiles across several online resources, supplied with and support multiple assessments to avoid any discrimination regarding genders or age groups in a national basis of the country.. [6].

The study "Leveraging AI to Transform Talent Acquisition" brings attention towards using technology i.e., implementation of artificial Intelligence (AI) and Machine Learning (ML) in optimizing the process of recruitment. It involves building predictive analytics models potentially, which can help in evaluating candidates on any dimension and sorting them accordingly as per qualification that may make the hiring process efficient & saves time for recruiting folks. In the end, the proposal argues a more inclusive hiring process which will be enabled by using AI and ML extensively for talent acquisition. [7].Paper titled Computer, Whom Should I Hire? How to Integrate AI Technology in the Process of Hiring Factors Influencing Its Acceptance Amongst Stakeholders acceptance criteria for Ai In Recruitment. Its objectives are to draw up AI Acceptance Criteria in Recruitment, prepare AI governance principles for best use and service optimisation trials before deployment. It references different kinds of AI systems being used in recruitment and how text analysis technologies add velocity to application processing. Sets of automated initial screens are illustrated in the evaluation of qualifications and soft skills by candidates. [8].

TABLE I. REPRESENTS THE LITERATURE ANALYSIS

Author/Year	Proposed Work	Techniques and Methods Used	Result
Jayant.ai(2022)	Improve recruitment screening by utilizing predictive analytics to anticipate candidate acceptance.	Implement machine learning algorithms such as Random Forest and Naïve Bayes;	Random Forest demonstrates superior performance over Naïve Bayes, achieving 71.5% accuracy and 83.4% recall.
Baykal.ai(2020)	Analyze current digital trends in recruitment, focusing on their advantages and disadvantages for organizations.	Review of Applicant Tracking Systems (ATS), online and social media recruitment strategies, automated screening processes, and data sharing practices.	E-recruitment enhances efficiency and cost-effectiveness;
Alrakhawi(2024)	Investigate the impact of AI on human resource management, specifically in recruitment, performance evaluation, and employee engagement.	Utilize AI-driven recruitment tools, predictive analytics, and conduct case studies while considering data privacy issues and the integration of blockchain and IoT technologies.	AI improves recruitment efficiency and personalizes the candidate experience.
Rhemananda.ai(2021)	Examine the role of blockchain technology in enhancing recruitment by ensuring better data validation and integrity.	Implement candidate registration on blockchain platforms, utilize automated data validation, and ensure data integrity through consensus mechanisms.	Improved data reliability and transparency in recruitment processes, especially advantageous for small and medium-sized enterprises (SMEs).
Allal-Chérif.ai(2021)	Assess AI's global influence on talent identification, selection, and retention.	Develop an intelligent recruitment system using big data analytics, grounded theory, and content analysis, focusing on minimizing bias related to gender, age, and nationality..	technologies streamline recruitment processes, reduce biases, and improve employee retention rates.
Trinh.ai(2021)	Create a natural language processing (NLP) framework for automating resume classification and ranking.	Utilize text preprocessing techniques, Named Entity Recognition (NER) with SpaCy, fastText word embeddings.	Enhanced processing of resumes and improved similarity assessments lead to better candidate ranking and greater efficiency in recruitment workflows.
Thakkar(2021)	Introduce AI-based predictive analytics tools to refine recruitment processes.	Apply predictive analytics, machine learning algorithms for candidate screening, debiasing techniques, and implement AI chatbots and virtual assistants.	Increased efficiency in recruitment, reduced workload for HR teams, and enhanced candidate engagement through AI-driven methodologies.
Laurim(2021)	Explore acceptance criteria for the integration of AI in recruitment from the perspective of stakeholders.	Conduct interviews with recruiters and candidates, utilize text analysis technologies, focus on automated screening, and categorize user control features and acceptance factors.	Effective integration of AI necessitates stakeholder engagement, clarity in AI functionalities.

TABLE II. REPRESENTS THE USE OF NEW EMERGING TECHNOLOGIES IN RECRUITMENT AND SELECTION PROCESS.[14]

Response	Frequency	Percent	Valid Percent	Cumulative Percent
Strongly Disagree	8	13.3%	13.3%	13.3%
Disagree	4	6.7%	6.7%	20.0%
Neutral	8	13.3%	13.3%	33.3%
Agree	4	6.7%	6.7%	40.0%
Strongly Agree	36	60.0%	60.0%	100.0%
Total	60	100%	100%	

III. EXISTING METHODS

A. Predictive Analytics and Machine Learning

- Predictive Analytics: This is evaluating historical data to make predictions about future hiring Predictive Analytics It helps to optimize recruitment processes, and identify suitable candidates analysing qualifications trends as well as cultural fit.
- Random Forest: This method produces a number of decision trees during the training phase and outputs the mode prediction. It improves the accuracy of a model while avoiding overfitting, and so is very useful against switch in problems.
- Naïve Bayes: it is a simply probabilistic model based on applying Bayesian theorem, considering features as independent of one another. It is fast and efficient, particularly useful in some tasks like resume filtering.

B. Data Management

- Data Collection: which consists on collecting recruitment data, specially CVs from the last 4 years to analyse some characteristics and candidate success rates afterwards we just need to see what profile it takes for a good outcome.
- Data Preprocessing: This step checks the cleanliness of dataset by removing duplicates, fixing errors and making homogeneous formats which increases data quality so that biases are reduced.
- Model Development: the platform consists in training predictive models on labelled data to allow for learning from past hires and predicting future candidate success.

C. E-Recruitment Practices

- Applicant Tracking Systems (ATS): The pick & and drop style interfaces which gather the right data without wasting time of HROPIn8 Deep analytics helping in Talent Insights Recruit at a faster phase with Video Interview Applicant Tracking Systems (ATS) to manage end to end recruitment under one platform for easy job posting, no more paper-based applications.
- Online Recruitment: here recruiters will have to use the internet to widen efforts on recruiting through job boards, company websites and even social media platforms making it more publicly visible of available openings.
- Social Media Recruitment: This method engages candidates with a presence on sites like LinkedIn and Facebook, allowing for specific engagement efforts as well web-intelligence based off their virtual profile.

D. AI Integration in HRM

- AI-Enhanced Recruitment Tools: These tools use AI algorithms to help you source candidates and automate the initial screening of applications against job requirements, so that you are able to evaluate large number of applicants in a non-biased way saving human effort.
- Data-Driven Performance Management: With continuous data capture on employee performance, organizations have a sound way of spotting trends and creating customized development agendas for growth specific to each individual.
- Integration with Other Technologies: AIOps can also be integrated with Blockchain which will provide better HR solutions and data security along with transparency.

TABLE III. REPRESENTS THE ADVANTAGES AND DISADVANTAGES OF TECHNIQUES

Category	Advantages	Disadvantages
Predictive Analytics and Machine Learning	Improves recruitment by forecasting outcomes, identifying ideal candidates, and reducing the chance of model overfitting.	Susceptible to data bias, can be challenging to select features, and often needs large data volumes.

Data Management	Improves data reliability, enhances model predictions, and offers insights into candidate success trends.	Requires extensive time for data cleaning, can carry bias if preprocessing isn't thorough, and demands high computing power.
E-Recruitment Practices	Accelerates hiring processes, extends candidate reach, allows targeted hiring, and minimizes paperwork.	Potential for bias, raises privacy issues, can result in overwhelming applications, and poses security challenges.
AI Integration in HRM	Increases hiring efficiency, helps reduce biases, and enables tailored employee development.	Risks bias in algorithms, raises privacy concerns, and often comes with significant setup costs.
Blockchain Technology in Recruitment	Protects data integrity, prevents fraudulent information, ensures data reliability, and simplifies verification.	High initial expenses, implementation complexity, and relies on broad adoption.
Research Techniques	Offers practical insights, uncovers trends, and aids in making well-informed recruitment choices.	Requires considerable resources, can be subjective, and presents privacy concerns.

IV. METHODOLOGY

A. Research Design

This research is on the automated Aquantitative research methodology is applied in this study with measurable outcomes like the level of resume to job profile alignment. This system aims to:

- Highlight specific skills or qualifications that are missing from resumes.
- Assign a compatibility score to candidates and job profiles.
- Create personalized recommendations for upskilling to mitigate skill shortages.

The framework is flexible and can be applied to a wide variety of industries and jobs.

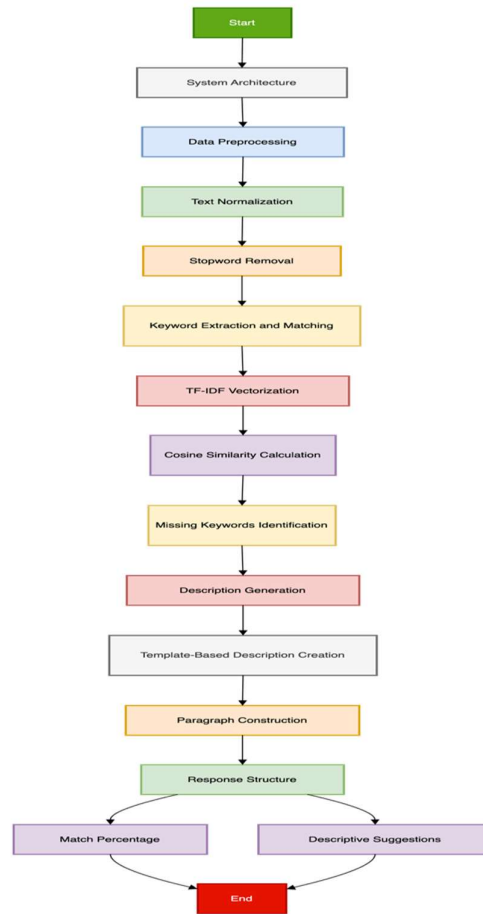


Figure 2. Represents the process

B. Data Collection

The model was developed and validated using a large-scale data set. Key aspects include:

The data is composed of resumes and job descriptions from public job boards, HR databases, and simulated datasets that syntactically resemble real-world recruitment situations.

Dataset Composition: Data included resumes in .docx format and job descriptions with required skills and qualifications and responsibilities. These datasets included a diversity of roles and levels of experience.

Ethical Safeguards: In accordance with GDPR and similar regulations, all personal information was anonymized. Ethical consideration was given for data use, and when needed, the consent was taken from participants.

C. Data Preparation

The data was pre-processed and the following steps were performed:

Normalization: The text was normalized by putting it all to lower case, removing digits, and stripping special characters for uniformity.

- **Stopword Filtering:** Library of common words with less global significance (e.g., “and,” “the”) were filtered out from text to increase relevancy of extracted terms.
- **Cleaning and Tokenization:** Resumes and job descriptions were broken down into their constituent words and irrelevant or redundant terms were eliminated for the sake of clarity and normalization.

This step made sure the dataset was clean and ready to be analyzed through ML and NLP techniques.

D. Feature Extraction

After preprocessing, textual data was converted into numerical formats to facilitate analysis:

- **Text Vectorization:** Weights for words were assigned using the TF-IDF (Term Frequency-Inverse Document Frequency) method to observe their importance in the dataset.
- **Similarity Measurement:** Cosine similarity was used to measure the alignment between the resumes and job descriptions. This score, written as a percentage, indicates the degree of match between a candidate’s profile and the job requirements.
- **Model Design and Evaluation:** The system utilizes Natural Language Processing (NLP) and Machine Learning (ML) techniques to process and score candidates. Key elements include:

Algorithm Integration

Tools like spaCy or NLTK were used for parsing and data extraction. Candidate ranking and classification were evaluated using ML models such as Random Forest and Naïve Bayes.

E. Personalized Recommendations

The system gives customized suggestions to improve resumes:

Recommendation Templates

Missing keywords for technical skills, soft skills, and project experience are filled in using predefined templates. Here are some recommendations to help candidates better mentor and develop themselves towards alignment their profiles with job requirements.

The screenshot displays the FastAPI web interface for a REST API. At the top, the 'FastAPI' logo is visible with version '0.70.0' and 'GAS 1.1'. Below the logo, the URL 'https://api.joon' is shown. The main interface is titled 'default' and shows the endpoint 'POST /match_resume/ Match Resume'. Under the 'Parameters' tab, it states 'No parameters'. The 'Request body' section is set to 'multipart/form-data'. There are two input fields: 'file' (required, string, file) with a 'Choose file' button and 'resume' text, and 'job_description' (required, string) with a text area containing 'Project Quality: Enterprise level, high-performance projects'. A large blue 'Execute' button is positioned below the inputs. The 'Responses' section shows a '200 Successful Response' with a 'Multi type' dropdown menu set to 'application/json'. Below the dropdown, it says 'Content: Accept header' and 'Response: Value: / Success'.

Figure 3. Represents the FastAPI page of file upload and job discription

Outcome

Candidates get actionable feedback on their profiles and have the opportunity to improve and increase their chances of selected.

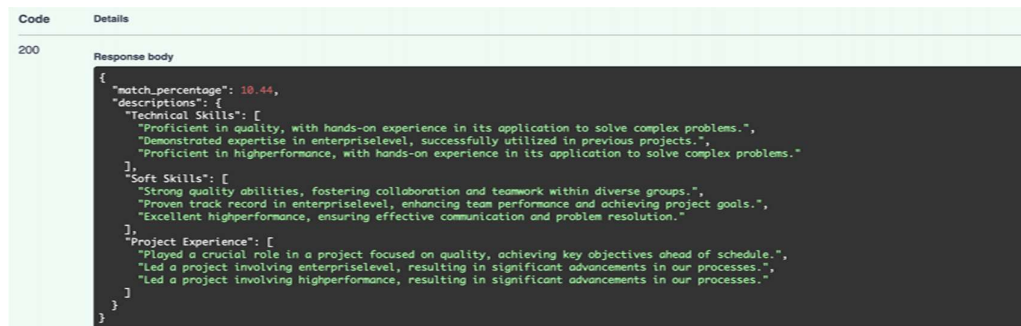
G. System Deployment

The deployed solution is hosted as a RESTful API with FastAPI, so it can be simply integrated into an existing HR tool:

Functionality: The API takes resumes and job descriptions as input.

Outputs include:

- A compatibility score.
- A list of missing skills.
- Personalized tips for improving your profile.



```
Code    Details
200     Response body
{
  "match_percentage": 10.44,
  "descriptions": {
    "Technical Skills": [
      "Proficient in quality, with hands-on experience in its application to solve complex problems.",
      "Demonstrated expertise in enterpriselevel, successfully utilized in previous projects.",
      "Proficient in highperformance, with hands-on experience in its application to solve complex problems."
    ],
    "Soft Skills": [
      "Strong quality abilities, fostering collaboration and teamwork within diverse groups.",
      "Proven track record in enterpriselevel, enhancing team performance and achieving project goals.",
      "Excellent highperformance, ensuring effective communication and problem resolution."
    ],
    "Project Experience": [
      "Played a crucial role in a project focused on quality, achieving key objectives ahead of schedule.",
      "Led a project involving enterpriselevel, resulting in significant advancements in our processes.",
      "Led a project involving highperformance, resulting in significant advancements in our processes."
    ]
  }
}
```

Figure 4. Represents the suggestion for improving the resume.

Integration

The API integrates with ATS (Applicant Tracking systems), which automates the candidate screening process, thus unburdening recruiters.

H. Ethical Compliance

In the interest of fairness and transparency, we took the following steps:

Bias Mitigation

The system assesses candidates purely on skills and experience and does not take demographic variables like gender, ethnicity or age into account.

Data Privacy

All processing of the data, in accordance with international data protection laws, ensures candidate confidentiality and security.

IV. CONCLUSION

In this paper, a complete framework is introduced to improve resume screening process with the aid of Natural Language Processing (NLP) and Machine Learning(ML). We hope that our resume evaluation automation will cut through inefficiencies and biases associated with traditional hiring processes. It combines powerful data management techniques, predictive analytics and ethical considerations to support an environment in which recruitment decisions are based on job-relevant qualifications rather than subjective biases. Our results indicate that it is possible to speed up resume screening using NLP and ML, yet reducing the impact of endemic biases towards fairness. By extracting features, systematic candidate ranking ensures that evaluations are based on the skills and experiences of candidates—and not (only) their names—which supports a more equitable hiring environment in turn. But they said this process should be used in an ongoing way to help keep ethical standards and understand how the labor market is changing.

FUTURE SCOPE

Points which are extracted Data Organization System This is the backbone of any resume parser. Implementing NLP models to process unstructured text present in resumes and extracting job history, skills sets, educational

background and certifications. It uses keyword and phrase matching methods to make sure a candidate has pretty much the experience needed by job condition. Named Entity Recognition (NER) Tags important entities such as company names, roles and locations. The system can also skillfully perceive different types (PDF, Word document) making it a seasoned expert at manipulating resumes with various structures.

Scoring and Ranking Candidates A targeted resume-level scoring system based on the weighted assessment of core attributes in each job such as skills, experience level, certifications, education etc. Based on that, ML models rate importance of candidate experience and skills to the JOB description output a data-driven Match score. A completeness check will flag any resumes that are missing the essential data again so recruiters have full profiles to work with. **Bias Detection and Reduction** is The solution also uses NLP-based modules and text formatting or linguistic patterns which, when applied to the descriptive information for these data subsets can potentially relax demographic biases. On their end, efforts include disincentives (omission of name and demographic details in processing) to prevent bias-based decision-making. Also, the organizations can use fair degree criteria such as diversity and inclusion that helps them in aligning with organizational equity goals.

REFERENCES

- [1] Jayanti, L. P. S. D., and Meditya Wasesa. "Application of Predictive Analytics To Improve The Hiring Process In A Telecommunications Company." *Jurnal CoreIT* 8, no. 1 (2022): 32-39.
- [2] Baykal, Elif. "Digital era and new methods for employee recruitment." In *Handbook of research on strategic fit and design in business ecosystems*, pp. 412-430. IGI Global, 2020.
- [3] Alrakhawi, Hazem AS, Randa Elqassas, Mohammed M. Elsobeihi, Basel Habil, Basem S. Abunasser, and Samy S. Abu-Naser. "Transforming Human Resource Management: The Impact of Artificial Intelligence on Recruitment and Beyond." (2024).
- [4] Rhemananda, Happy, Dima Roulina Simbolon, and Olivia Fachrunnisa. "Blockchain technology to support employee recruitment and selection in industrial revolution 4.0." In *Proceedings of International Conference on Smart Computing and Cyber Security: Strategic Foresight, Security Challenges and Innovation (SMARTCYBER 2020)*, pp. 305-311. Springer Singapore, 2021.
- [5] Allal-Chérif, Oihab, Alba Yela Aránega, and Rafael Castaño Sánchez. "Intelligent recruitment: How to identify, select, and retain talents from around the world using artificial intelligence." *Technological Forecasting and Social Change* 169 (2021): 120822.
- [6] Trinh, Quynh, and Thanh-Tuan Dang. "Automatic process resume in talent pool by applying natural language processing." In *Proceedings of international conference on logistics and industrial engineering 2021*, pp. 234-240. 2021.
- [7] Thakkar, Darshit. "Leveraging AI to transform talent acquisition." *International Journal of Artificial Intelligence and Machine Learning* 3, no. 3 (2021): 7.
- [8] Laurim, Vanessa, Selin Arpacı, Barbara Prommegger, and Helmut Kremer. "Computer, whom should i hire?—acceptance criteria for artificial intelligence in the recruitment process." (2021).
- [9] Singh, A. N. U. K. R. I. T. I., and D. M. K. Sahoo. "Revolutionizing recruitment: harnessing the power of technology." *IRE J* 6 (2023): 1343-1355.
- [10] El-Ghoul, Mostafa, Mohammed M. Almassri, Mohammed F. El-Habibi, Mohanad H. Al-Qadi, Alaa Abou Eloun, Bassem S. Abu-Nasser, and Samy S. Abu-Naser. "AI in HRM: Revolutionizing Recruitment, Performance Management, and Employee Engagement." (2024).
- [11] Lakshan, W. G. K., K. A. V. Prabuddhi, P. P. Weerasinghe, E. M. L. P. Bandara, M. P. Gamage, and R. R. P. De Zoysa. "Revolutionizing the Hiring Process with Automated Evaluation and Behavioral Analysis–IntelliHire." *International Research Journal of Innovations in Engineering and Technology* 7, no. 11 (2023): 307.
- [12] Brishti, Juthika Kabir, and Ayesha Javed. "The viability of ai-based recruitment process: A systematic literature review." (2020).
- [13] Hemalatha, A., P. Barani Kumari, Nishad Nawaz, and Vijayakumar Gajenderan. "Impact of artificial intelligence on recruitment and selection of information technology companies." In *2021 international conference on artificial intelligence and smart systems (ICAIS)*, pp. 60-66. IEEE, 2021.
- [14] Prabhakar, Neve Jayant. "Role of new emerging technologies in recruitment and selection process." *The Online Journal of Distance Education and e-Learning* 11, no. 1 (2023).