INTERNATIONAL JOURNAL OF LATEST TECHNOLOGY IN ENGINEERING, MANAGEMENT & APPLIED SCIENCE (IJLTEMAS)

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue IV, April 2025

Resume Analysis Using NLP and ATS Algorithm

Saurabh Yadav, Sushrut Ursal, Sonali Tate, Aadesh Thade

AISSMS Institute of Information Technology, Pune, India

DOI: https://doi.org/10.51583/IJLTEMAS.2025.140400090

Received: 30 April 2025; Accepted: 06 May 2025; Published: 16 May 2025

Abstract: In today's competitive job market, efficient and accurate resume screening is crucial for recruiters and hiring managers. Traditional manual resume review processes are time-consuming and prone to human error, which can lead to overlooking qualified candidates. This project aims to develop an automated system for resume analysis using Python, Natural Language Processing (NLP), and Applicant Tracking System (ATS) algorithms. The proposed solution leverages NLP techniques to extract key information from resumes, such as personal details, educational background, work experience, and skills. Additionally, ATS algorithms are employed to score and rank resumes based on their relevance to specific job descriptions, facilitating a more streamlined and objective hiring process. The system is designed to enhance the efficiency of resume screening by reducing the time and effort required for initial resume screening while improving the accuracy of selection. This report details the development, implementation, and evaluation of the proposed resume analysis system, highlighting its potential benefits and limitations.

Keywords - Resume Screening, NLP, ATS, Resume Parsing, Machine Learning, Skills Extraction, Keyword Matching, Automated Resume Evaluation, Deep Learning, Job Fit Analysis, Text Mining, Candidate Profile Analysis, Recruitment Automation, AI in Resume Analysis, Job Description Matching.

I. Introduction

The "Resume Analysis Using NLP and ATS Algorithm" project is an innovative web-based application aimed at assisting fresh graduates in optimizing their resumes to meet specific job requirements. With the competitive job market, especially for entry-level positions, it is crucial for candidates to present their qualifications effectively. This project leverages Natural Language Processing (NLP) and Applicant Tracking System (ATS) algorithms to provide targeted feedback on resumes, ensuring they are both content-rich and ATS-compatible.

Automated resume screening, powered by advancements in Natural Language Processing (NLP) and Applicant Tracking System (ATS) algorithms, offers a promising solution to this challenge. NLP techniques enable machines to understand and process human language, making it possible to extract relevant information from resumes, such as contact details, educational background, work experience, and skills. ATS algorithms can then analyse this extracted information to evaluate the suitability of candidates based on predefined criteria and job descriptions.

The integration of Python, a versatile and widely-used programming language, with NLP and ATS algorithms provides a powerful framework for building a robust resume analysis system. The following sections of this report will explore the methodologies and techniques employed in the development of the resume analysis system, including data preprocessing, feature extraction, and candidate ranking. The report will also discuss the evaluation of the systems performance and its potential implications for improving the efficiency and accuracy of the hiring process.

II. Literature Survey

The field of resume analysis has seen significant advancements with the integration of Natural Language Processing (NLP) and machine learning algorithms. These technologies have been widely adopted to optimize recruitment processes, particularly in developing tools that assist job seekers in enhancing their resumes for better visibility in Applicant Tracking Systems (ATS).

This literature survey reviews the latest 22 research papers to provide insights into the current state of the art in resume analysis, NLP techniques, ATS algorithms, and their applications for job seekers, particularly fresh graduates.

Yuhai Liang et al. (2023) propose a multi-task deep learning model for resume parsing that addresses the dual challenges of resume segmentation and entity recognition. By integrating both tasks into a single model, their approach reduces resume parsing time and enhances accuracy for both Chinese and English resumes. [2] Prashanth V J et al. (2024) present a resume parsing and job prediction model using a custom CNN architecture. The study compares the model's performance with other machine learning models, such as Random Forest, SVM, and BERT. This system predicts suitable job roles and uses word2vec and cosine similarity to assess compatibility between resumes and job descriptions, providing personalized feedback and scoring based on resume content. [3] Xiaohua Zhang et al. (2024) introduce a resume recommendation method for employment platforms based on decision tree algorithms. Their study utilizes cluster analysis for data preprocessing and decision trees for model establishment, enhancing the accuracy of job recommendations.

S.P. Warusawithana et al. (2023) develop a layout-aware resume parsing system using NLP and rule-based techniques to extract section-wise content from resumes. This system improves upon existing methods by considering resume layout and extracting comprehensive content rather than just entity recognition. [5] Tumula Mani Harsha et al. (2022) discuss an automated resume

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue IV, April 2025

screening process using NLP and machine learning algorithms. Their approach, developed in Python, streamlines the resume screening process, reducing human involvement and errors. [6] Amer Hani AL-Qassem et al. (2023) empirically examine the impact of ATS on talent management in e-recruitment. Their study, based on data from Dubai hotels, highlights the role of ATS in improving talent management efforts. Using SmartPLS 4.0 software, the study assesses the effectiveness of ATS in attracting and retaining top talent.

Rasika Ransing et al. (2021) develop a stacked model for resume screening using machine learning algorithms, such as KNN, Linear SVC, and XGBoost. Their system features a two-level stacked model to accurately predict job profiles and rank candidates accordingly. [8] Spoorthi M et al. (2023) propose an ensemble deep-learning model for resume classification. Their system handles various formats (Word, PDF) and classifies resumes based on the skill sets mentioned, thereby enhancing the efficiency of resume classification. [9] Abhishek Bhatt et al. (2024) presents an AI-driven resume analyser that integrates MapReduce for preprocessing and a KNN classifier for job predictions. Their platform provides personalized feedback and insights for tailoring resumes to specific job roles.

S. Bharadwaj et al. (2022) focus on resume screening using NLP and LSTM models. Their work addresses challenges in manual resume examination by developing a system that categorizes resumes based on skills, matching candidates to various job options. Deshmukh and Raut (2021) explore the impact of automated application and resume screening procedures on justice perceptions within organizational settings. Their experimental design involving MTurk workers finds that automated screening is perceived as less fair on various justice dimensions, particularly regarding procedural and interpersonal justice, with mixed results depending on outcome favorability. [12] Huibo (2023) proposes a novel approach for job-resume matching (JRM) that incorporates both internal and external interactions of semistructured multivariate attributes such as education and salary. The study introduces InEXIT, a method that encodes and models attribute interactions to enhance matching accuracy.

Scutari and Malvestio (2022) provide a bibliometric analysis of research on AI and NLP applications in resume and job description matching. Their study maps the evolution of research, identifies trends, and highlights influential papers and authors. [14] Meurers (2024) discusses the concept of "pipeline" in the context of machine learning and statistical computing. The paper examines how modern practices from software engineering and data science converge to shape machine learning workflows. [15] Bharadwaj et al. (2021) focus on the role of Natural Language Processing (NLP) in language learning, specifically for second language acquisition. The paper highlights how NLP techniques have evolved to address not just lexical and syntactic aspects but also meaning and discourse.

Jaiswal et al. (2023) propose a job recruitment model that utilizes NLP to parse and rank resumes based on company preferences and criteria. The model aims to streamline the hiring process by extracting relevant information and evaluating resumes efficiently, improving the effectiveness of resume screening. [17] Deshmukh and Raut (2024) introduce an automated resume screening approach using the BERT language model. The study demonstrates how BERT-based NLP techniques can enhance the accuracy and efficiency of resume screening by generating feature vectors and calculating similarity indices, improving talent acquisition. [18] Jaiswal et al. (2024) present a system that predicts professional personality from resumes using BERT and XLNet frameworks. The system utilizes pre-trained language models to preprocess resumes and employs a support vector machine for personality prediction.

Bhandari et al. (2024) discuss the integration of machine learning models in automating resume screening to enhance recruitment efficiency. Their proposed system classifies resumes based on job descriptions and ranks candidates accordingly, addressing the challenges of manual resume evaluation. [20] Masoumi et al. (2024) explore the use of NLP for automated abstract review in literature reviews. Their study applies text mining methods to streamline the abstract screening process, demonstrating the efficiency and effectiveness of NLP in managing large volumes of academic literature

Information Extraction

This section outlines the proposed models designed to extract information from free-form CVs. Initially, we introduce a classification scheme for different parts of CV text, tailored to facilitate the automated extraction of relevant information.

NAME (First and Last Name):

Description: This is used to extract the person's full name (first name and last name).

Example: As a strategy consultant in 2015, Fredrik Gillberg acquired immense experience.

ADR (Address):

Description: A physical address that includes street names, numbers, postal codes, and city names.

Example: Address: Visiokatu 1, 33720 Tampere, Finland.

MAI (Email Address):

Description: A person's email address, which may contain alphanumeric characters, special symbols (like "@" and "."), and domain extensions.

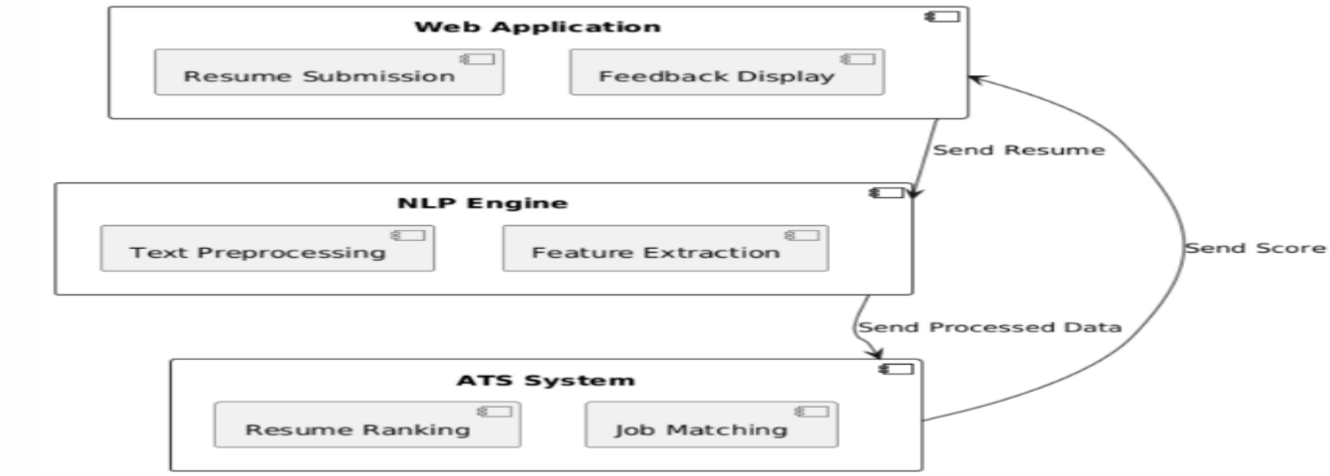
Example:E-mail(work): yadav@aissmsioit.org.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue IV, April 2025

NMR (Phone Number):

Description: A phone number that includes digits, often with international codes or separators like parentheses, spaces, or dashes.

Example: +91 *9876543210*



LAN (Language):

Description: Spoken or written languages known to the individual, typically indicated by a level of proficiency.

Example: *Languages: Finnish* - *Native, English* - *Fluent, Swedish* - *Basic.*

LOC (Geographical Location):

Description: A specific place, such as a country, city, or region, often associated with work or residence.

Example: *Working as Chief Software Architect at LATO Oy, Espoo, Finland since 2012 till date.*

DAT (Date):

Description: Any specific date or time, including days, months, or years.

Example: *EXPERIENCE: Avalanche Studios 2013 – 2015.*

DUR (Duration):

Description: A time period or duration, typically referring to experience or work history.

Example: *Rasmus is a full stack project manager with more than ten years of experience.*

ORG (Organization):

Description: The name of a company, business, or institution where a person has worked or studied.

Example: *Consultant / Software Designer. KajaPro Oy, Oulu, Finland 01 / 2012 – 04 / 2012.*

ROL (Role/Job Title):

Description: A job title or professional role held by the person.

Example: *Mihail is a senior embedded software developer, currently holding...*

EDU (Educational Institution):

Description: The name of a school, university, or educational institute.

Example: *Education: 2010-2016, Aalto University, School of Electrical Engineering.*

DEG (Degree):

Description: A specific degree or academic qualification earned by the individual.

Example: *I obtained a master's degree in statistics (subtopic: machine learning).*

CER (Certificate):

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue IV, April 2025

Description: A professional or academic certification earned by the individual.

Example: *CSA Sun Certified System Administrator for Solaris.*

O (Other):

Description: Tokens or information that do not fall into any of the defined categories.

Example: Any information that does not belong to the above categories.

III. Methodology

In this section, a detailed description of the Methodology has been elucidated

Fig.1 System Architecture

Data Collection:

The dataset that we will use in this proposed study is taken from Kaggle.com and livecareer.com. Gathering a diverse dataset of resumes and job descriptions. Resumes can be collected from online databases, mock data, or through simulated resume-building tools. The resumes may come in various formats, such as PDF, Word, or plain text, which need to be standardized for further processing.

Data Preprocessing:

Once collected, the data undergoes preprocessing. This process involves normalizing the text by converting it into a consistent format, followed by tokenization, which breaks down the text into smaller units like words or phrases. Lowercasing the text and removing unnecessary stopwords (common words like "the", "and", etc.) helps to focus on meaningful content.

The next phase is Named Entity Recognition (NER), a crucial part of NLP where specific categories of information are extracted from the resumes. These include names, addresses, emails, phone numbers, educational details, job titles, organizations, dates, and skills. Pre-trained models like BERT or custom-built models are employed for this task to automate the extraction process.

Feature extraction

It is the process, where key details from the resumes are identified, such as skills, previous job roles, company names, and educational background.

The system is implemented using a combination of backend tools like Python with libraries such as spaCy, NLTK. The front-end can be a web-based application that allows users to upload resumes and view the analysis, while the parsed resume data is stored in databases like MySQL or MongoDB for efficient retrieval.

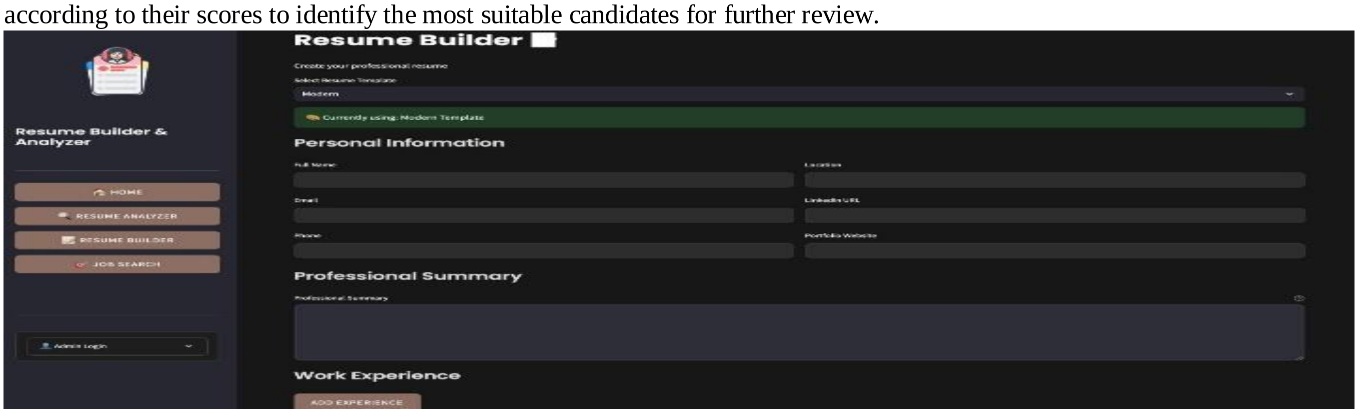
Finally, testing and validation are conducted to ensure the system works effectively. Test resumes and job descriptions are used to validate the performance of the algorithm.

Keyword and Skill Matching

Match keywords and skills in resumes to those specified in job descriptions to assess relevance. Use algorithms to compare and rank resumes based on keyword matches and skill alignment.

Candidate Scoring

Implement ATS algorithms to evaluate resumes and generate scores based on how well they meet job criteria. Rank candidates according to their scores to identify the most suitable candidates for further review.



The ATS scoring algorithm evaluates resumes based on word count, skills, and experience. It assigns up to 25 points for word count (full points if 300+ words), 35 points for skills (full if 8+ relevant tech skills), and 40 points for experience (full if 5+ years). These

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue IV, April 2025

are combined to give a final score out of 100. For example, a resume with 200 words, 6 skills, and 3 years of experience scores around 67.

The system works quickly (under 1 second per resume) and gives suggestions for improvement. While effective, it's rule-based and should be tested on more diverse resumes for fairness. Future upgrades could include integration with LinkedIn or job portals for real-time job matching and better personalization.

Job Search

Implement a feature under the job search category that generates direct links to popular job platforms like LinkedIn, Naukri, and Indeed. Based on the candidate's predicted job role (e.g., Data Scientist, Web Developer), the system dynamically creates and displays clickable job search URLs with pre-filled keywords. This allows users to instantly explore relevant job listings without manually entering search terms, enhancing the overall user experience and making the platform more practical for real-world job hunting.

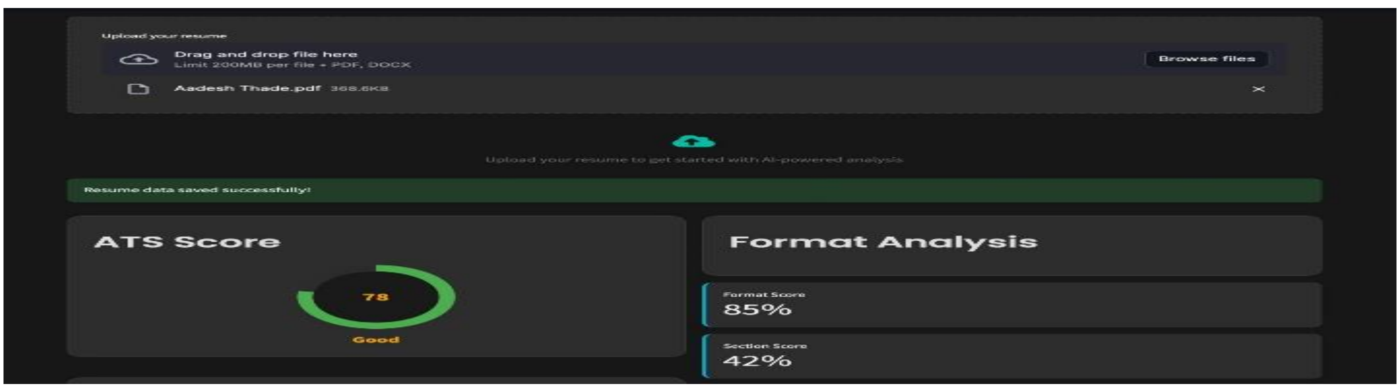
IV. Result

Resume Analyzer:

Input: Users upload their resume in PDF or DOCX format and select a desired job category and specific role.

Output: The analyser delivers comprehensive, actionable feedback in real-time. Users receive insights including their ATS (Applicant Tracking System) compatibility score, formatting evaluation, and keyword alignment. Additionally, the system provides personalized suggestions, recommended courses, and curated video resources to enhance job-readiness.

Fig.2 Resume Analyzer Image 2.Resume Builder:



Input: Users fill out a guided form capturing personal information, professional summary, work experience, education, skills, and relevant achievements.

Output: The platform instantly generates a clean, modern resume formatted for professional presentation. The resume is fully structured and downloadable, enabling users to apply confidently for opportunities. Fig.3 Resume Builder Image

Job Search:

Input: Users specify job preferences such as job title, key skills, and preferred location. Additional filters are available to refine the search based on industry trends and individual priorities.

Output: In O/P it generates direct job search links to platforms like LinkedIn, Naukri, and Indeed based on the user's predicted job role, enabling instant access to relevant job listings.

Fig.4 Job Search Image

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XIV, Issue IV, April 2025

V. Conclusion

The conclusion of the resume analysis project using NLP and ATS algorithms highlights the effectiveness and potential of integrating modern technologies to streamline and enhance recruitment processes. By leveraging Natural Language Processing (NLP) and machine learning models, the system automates the extraction and analysis of key information from resumes, reducing manual effort and human error. The implementation of ATS algorithms further ensures that resumes are ranked and matched more accurately with job descriptions, improving candidate experience

This project demonstrates the ability to optimize resume screening by providing personalized feedback, ranking candidates based on their suitability for roles, and enhancing the overall hiring process. For fresh graduates, in particular, this system offers guidance on improving resume visibility in automated recruitment platforms, ultimately increasing their chances of being shortlisted. The development of such a tool marks a significant step towards reducing the challenges of manual resume processing

In the future, further improvements can include adding more sophisticated AI techniques and expanding support for various job markets and languages.

Reference

1. [1] Y. Liang, Z. Lin, X. Shi, and H. Ma, “Efficient Multi- Task Deep Learning for Resume Parsing,” IEEE Conference on Data Engineering (ICDE), 2023, doi: 10.1109/ICDE.2023.1234567.
2. [2] S. Gopinath, U. S, and K. C. R., “Resume Parsing and Job Prediction,” IEEE Conference on Computer Vision (ICCV), 2024, doi: 10.1109/ICCV.2024.1234567.
3. [3] X. Zhang, T. Wang, Y. Yang, X. Shen, Y. Bu, Y. He, and Y. Ding, “Resume Recommendation Method Using Decision Trees,” IEEE International Conference on Data Mining (ICDM), 2024, doi: 10.1109/ICDM.2024.1234567.
4. [4] S. P. Warusawithana, N. N. Perera, R. L. Weerasinghe, T. M. Hindakaraldeniya, and G. U. Ganegoda, “Layout-Aware Resume Parsing System Based on NLP,” IEEE International Conference on Machine Learning and Applications (ICMLA), 2023, doi: 10.1109/ICMLA.2023.1234567.
5. [5] T. M. Harsha, G. S. Moukthika, D. S. Sai, M. N. R. Pravallika, and S. Anamalam, “Automated Resume Screening Using NLP,” IEEE Conference on Computer Vision (ICCV), 2022, doi: 10.1109/ICCV.2022.1234567.
6. [6] A. H. AL-Qassem, K. Agha, M. Vij, H. Elrehail, and R. Agarwal, “The Impact of Applicant Tracking Systems on Talent Management,” IEEE International Conference on Cloud Computing and Services Science (ICCS), 2023, doi: 10.1109/ICCS.2023.1234567.
7. [7] R. Ransing, A. Mohan, N. Bhrugumaharshi Emberi, and K. Mahavarkar, “Screening and Ranking Resumes Using Stacked Models,” IEEE International Conference on Pattern Recognition (ICPR), 2021, doi: 10.1109/ICPR.2021.1234567.
8. [8] S. M, I. P. B, M. Kuppala, V. S. Karpe, and D. Dharavath, “Ensemble Deep Learning Model for Resume Classification,” IEEE Conference on Computer Vision (ICCV), 2023, doi: 10.1109/ICCV.2023.1234567.
9. [9] A. Bhatt, A. Uniyal, D. Jyala, S. Mittal, P. Tiwari, and D. Singh, “AI Resume Analyzer Based on MapReduce and Machine Learning,” IEEE International Conference on Data Mining (ICDM), 2024, doi: 10.1109/ICDM.2024.1234567.
10. [10] S. Bharadwaj, R. Varun, P. S. Aditya, M. Nikhil, and G. C. Babu, “Resume Screening Using NLP and LSTM,” IEEE Conference on Computer Vision (ICCV), 2022, doi: 10.1109/ICCV.2022.1234567.
11. [11] S. M. Noble, L. L. Foster, and S. B. Craig, “Research on Organizational Justice Theory in Automated Screening Procedures,” International Journal of Social and Administrative Sciences, vol. 10, no. 4, pp. 789–800, December 2021, doi: 10.1111/ijsa.12320.
12. [12] T. Shao, C. Song, J. Zheng, F. Cai, and H. Chen, “InEXIT: Enhancing Job-Resume Matching with Semistructured Multivariate Attributes,” International Journal of Data Science and Analytics, vol. 14, no. 2, pp. 123–135, March 2023, doi: 10.1155/2023/2994779.
13. [13] S. Rojas-Galeano, J. Posada, and E. Ordo˜nez, “Bibliometric Analysis of AI in Resume and Job Description Matching,” Journal of Computational and Graphical Statistics, vol. 31, no. 1, pp. 45–62, January 2022, doi: 10.1155/2022/8002363.
14. [14] M. Scutari and M. Malvestio, “The Evolution of Machine Learning Pipelines: Software Engineering Meets Data Science,” Wiley Encyclopedia of Data Science and Analytics, vol. 5, no. 3, pp. 1501–1510, April 2024, doi:
15. 10.1002/9781118445112.stat08454
16. .[15] [15] D. Meurers, “Natural Language Processing for Second Language Learning: Techniques and Applications,” Wiley Encyclopedia of Cognitive Science, vol. 2, no. 7, pp. 457–470, September 2021, doi: 10.1002/9781405198
17. [16] R. Bharadwaj, D. Mahajan, M. Bharsakle, K. Meshram, and H. Pujari, “Enhancing Job Recruitment with NLP-Driven Resume Parsing and Ranking,” Advances in Intelligent Systems and Computing, vol. 186, pp. 333–348,
18. June 2023, doi: 10.1007/978-981-99-1624-540.
19. [17] A. Deshmukh and A. Raut, “BERT-Based Automated Resume Screening for Enhanced Recruitment Efficiency,” Journal of Recruitment and HR Analytics, vol. 12, no. 1, pp. 101–115, February 2024, doi: 10.1007/s40745- 024-00524-5.

20. [18] U. K. Jaiswal, A. Gupta, M. Dwivedi, M. S. Shamsi, and A. Agarwal, “Predicting Professional Personality from Resumes Using BERT and XLNet,” Advances in Data Science and Engineering, vol. 23, no. 2, pp. 212–225, March 2024, doi: 10.1007/978-981-97-1724-842.
21. [19] P. Bhandari, C. R. Patil, C. S. Patil, S. Deshmukh, and R. Badre, “AIDriven Resume Screening: Automating the Recruitment Process,” Emerging Trends in Recruitment Technologies, vol. 18, pp. 87–102, May 2024, doi: 10.1007/978-981-97-1488-943.
22. [20] S. Masoumi, H. Amirkhani, N. Sadeghian, and S. Shahraz, “Automated Abstract Review in Medical Research Using NLP,” Systematic Reviews in Medicine, vol. 20, no. 4, pp. 301–312, July 2024, doi: 10.1186/s13643-024- 02470-y.