© May 2025 | IJIRT | Volume 11 Issue 12 | ISSN: 2349-6002

SMART ATS: Intelligent Resume Screening Using Machine Learning

Vivek M[1], Dhivya V[2], Rahul B Asangi[3], Sanketh L Reddy[4]
[1,2,3,4] *Computyer Science & Engineering Sir M Visvesvaraya Institute Of Technology, Bengaluru, India*

***Abstract*—This paper introduces SMART ATS, a machine learning-based system designed to automate resume screening. It uses natural language processing to extract and analyze key details from resumes, matching them with job descriptions to generate relevance scores. The system streamlines recruitment by reducing manual effort, improving candidate shortlisting, and promoting efficient, unbiased hiring.**

***Index Terms*—Applicant Tracking System (ATS), intelligent resume screening, machine learning, natural language processing (NLP), resume parsing, job-candidate matching, automated recruitment systems, candidate ranking, feature extraction, TF-IDF, support vector machines (SVM), Random Forest, Naive Bayes, bias mitigation in AI, ethical hiring, talent acquisition, recruitment analytics, semantic matching, artificial intelligence in human resources, resume classification.**

I. INTRODUCTION

Recruiters face growing challenges in manually screening large volumes of resumes, often resulting in delays and inconsistent hiring decisions. Traditional Applicant Tracking Systems (ATS) rely on simple keyword matching, which may miss qualified candidates. This paper introduces SMART ATS, an intelligent resume screening system that uses machine learning and natural language processing to automatically extract, analyze, and match resume content with job descriptions. The system improves screening efficiency, reduces human effort, and supports fair, data-driven recruitment.

A. Background and Motivation

In today's competitive job market, recruiters are overwhelmed by the volume of resumes received for each job posting. Manual screening is time-consuming and often prone to bias or error, while traditional Applicant Tracking Systems (ATS) mostly rely on keyword matching, which lacks contextual understanding. These challenges create a need for intelligent systems that can streamline the recruitment process and improve the quality of hiring decisions.

B. Objective of the Study

This paper proposes SMART ATS, an automated resume screening system that leverages machine learning and natural language processing. The objective is to extract relevant candidate information, evaluate it against job requirements, and rank resumes based on suitability. The system aims to enhance hiring efficiency, reduce recruiter workload, and support objective and consistent decision-making.

II. LITERATURE SURVEY

The evolution of Applicant Tracking Systems (ATS) has been closely tied to advancements in Machine Learning (ML) and Natural Language Processing (NLP). Traditional ATS systems primarily utilized keyword-based matching algorithms to filter resumes. These early systems, however, were limited by their inability to understand the context of terms or assess the actual relevance of a candidate's experience. Smith et al. (2019), in their work *"Keyword Matching in Resume Screening: A Review of Limitations and Future Prospects,"* discussed the shortcomings of such keyword-based methods and proposed integrating semantic search to improve the effectiveness of ATS systems. Their findings highlighted the need for more intelligent systems capable of interpreting resumes beyond simple keyword occurrences.

As the limitations of traditional ATS became apparent, Machine Learning emerged as a powerful alternative. In their study, Patel et al. (2020) presented an ML-based resume screening model in *"Automated Resume Screening Using Natural Language Processing,"* where they utilized Naive Bayes classification to categorize resumes based on job descriptions. Their approach, which relied on feature extraction and TF-

IDF (Term Frequency-Inverse Document Frequency), showed significant improvements over traditional systems by capturing more nuanced aspects of resumes. Similarly, Singh and Mehta (2021) explored Support Vector Machines (SVM) for resume classification in their research *"An Intelligent Resume Screening System Using Support Vector Machines"*, which also demonstrated promising results in accurately classifying resumes based on relevant skills and qualifications.

The integration of deep learning techniques into resume screening systems further enhanced their accuracy and efficiency. Li and Zhang (2022) in *"Leveraging BERT for Resume Parsing and Screening"* explored the use of BERT (Bidirectional Encoder Representations from Transformers) for resume parsing. By fine-tuning this pre-trained model, they were able to significantly improve the system's ability to understand the semantic meaning of resumes. Their work outperformed traditional machine learning models in terms of precision and recall, marking a significant step forward in NLP-driven ATS systems. The ability of transformer-based models like BERT to understand context and semantic relationships in resumes was a key advancement in the field.

Despite the promising advancements in AI-driven resume screening, several challenges persist. A major concern highlighted by Kumar et al. (2023) in *"Ethical Considerations in AI-Based Recruitment Tools"* is the potential for bias in AI models, which can perpetuate existing inequalities in hiring practices. The authors emphasized the need for ethical considerations in the development of AI recruitment tools, particularly in ensuring that training data is diverse and free from biases. Furthermore, data privacy is a critical issue, as resumes often contain sensitive personal information. Compliance with regulations such as GDPR remains a significant challenge for organizations implementing AI-driven ATS systems. These concerns highlight the need for careful design and continuous monitoring of AI-powered resume screening systems to ensure fairness and privacy.

III. METHODOLOGY

The development of the SMART ATS Resume Analyzer follows a systematic approach involving data collection, feature extraction, model training, and integration into a user-friendly application. The methodology aims to ensure that the system can efficiently parse resumes, classify them based on relevance to job descriptions, and provide an unbiased assessment. The following steps outline the core processes involved in the development of the system.

A. Data Collection and Preprocessing

For the development of the SMART ATS Resume Analyzer, a comprehensive dataset of resumes was collected from various online job portals, including those tailored to different industries, such as technology, healthcare, and finance. The dataset includes resumes with varied formats, such as PDF, DOCX, and plain text, containing relevant fields such as personal details, educational background, skills, work experience, and achievements. Preprocessing of this data involved the extraction of relevant text from these documents using text extraction libraries like PyPDF2 for PDF files and python-docx for DOCX files. Subsequently, the data was cleaned by removing irrelevant information such as headers, footers, and any non-textual elements, ensuring that only relevant sections of the resume were retained. Additionally, tokenization and lemmatization were applied to convert the text into a uniform format, reducing variations in wording and improving the effectiveness of the subsequent machine learning models.

B. Feature Extraction and Model Training

In the next phase of development, relevant features were extracted from the resumes to represent the textual data in a way that machine learning algorithms could process. Key features included skills, experience, education, and certifications, which were identified using Named Entity Recognition (NER) and Part-of-Speech (POS) tagging. Additionally, TF-IDF (Term Frequency-Inverse Document Frequency) was utilized to represent the resumes as numerical vectors, highlighting the importance of terms within each resume in relation to the entire dataset. Once the features were extracted, the dataset was split into training and testing sets. Various machine learning algorithms, such as Naive Bayes, Support Vector Machines (SVM), and Random Forest, were trained on this data to classify resumes into categories based on their relevance to a specific job description. Hyperparameter tuning was performed using techniques like Grid Search to improve the models' performance, with accuracy, precision, recall, and F1-score as evaluation metrics.

C. Integration and Evaluation

The trained models were integrated into a web-based application designed for easy interaction by recruiters and hiring managers. This application allowed users to upload resumes in multiple formats, which were then automatically parsed and analyzed by the model. The results, including a relevance score for each resume based on the job description, were displayed in a user-friendly interface. To evaluate the effectiveness of the SMART ATS Resume Analyzer, a set of real-world job descriptions and corresponding resumes were used as test cases. The model's predictions were compared against human recruiters' assessments to assess the accuracy and fairness of the system. In addition, the system was tested for bias in resume screening by evaluating how the model performed across different demographic groups, ensuring that the AI did not favor certain candidates based on gender, ethnicity, or other biased factors. Regular feedback loops were incorporated to continuously improve the model's performance by retraining it with new data.

IV. FINDINGS

The findings from the SMART ATS Resume Analyzer research show that the system performs well in evaluating resumes and generating relevant feedback for job seekers and recruiters. The Random Forest model achieved the highest accuracy at 85%, followed by the Support Vector Machine (SVM) with an accuracy of 82%, and Naive Bayes at 76%. These models were evaluated based on their ability to match resumes to job descriptions, and Random Forest was found to provide the most accurate and consistent results.

Table 1 below summarizes the performance of the machine learning models used in the analysis, showing the metrics of accuracy, precision, recall, and F1-score:

Model	Accuracy	Precision	Recall	F1-Score
Random Forest	85%	84%	86%	85%
SVM	82%	80%	83%	81%
Naïve Bayes	76%	75%	77%	76%

In addition to evaluating model performance, the system was also tested for bias related to gender, ethnicity, and resume formatting. The bias analysis showed that the SMART ATS Resume Analyzer performs fairly across diverse demographic groups. However, the system exhibited minor challenges when processing resumes with unusual or non-standard formats, which can affect its prediction accuracy for certain resumes.

Table 2 presents the results of the system's performance across various resume formats to demonstrate its adaptability:

Resume Format	Accuracy	Comments
Standard	85%	Higher accuracy with clear structure
PDF	81%	Slight decrease due to formatting issues
Image	75%	Lower accuracy due to image to text conversion
Unconventional	76%	Challenges in parsing nonstandard resumes

These findings confirm that the SMART ATS Resume Analyzer is an effective tool for automating the recruitment process, delivering accurate assessments and feedback to both recruiters and job candidates. However, further improvements are needed, particularly in handling diverse resume formats to enhance the system's adaptability to various industries.

V. CONCLUSION

In conclusion, the Smart ATS Resume Analyzer successfully utilizes machine learning techniques to enhance the recruitment process by efficiently analyzing resumes and matching them with job descriptions. The system improves both the speed and accuracy of candidate selection, making it a valuable tool for HR professionals. By automating resume screening, it reduces human bias and errors, ultimately contributing to a more effective hiring process. The research demonstrates the potential of AI in streamlining recruitment and offers insights into future advancements in this domain. Further enhancements could include better integration with other HR tools for seamless workflow.

VI. ACKNOWLEDGEMENT

I would like to express my heartfelt gratitude to my project guide, Dr. Dhivya V, Assistant Professor, for her exceptional guidance, unwavering support, and valuable insights throughout the course of this project. Her expertise and encouragement have been instrumental in shaping my research and helping me overcome various challenges. I would also like to extend my sincere thanks to Dr. Anitha **T N**, Head of the Department of Computer Science and Engineering, for her continuous support, resources, and motivation, which greatly contributed to the success of this project. I am deeply grateful to my friends and family for their constant encouragement, understanding, and belief in me, which kept me motivated during difficult times. Their unwavering support has been a source of strength, and I truly appreciate their patience and understanding throughout the research journey.

REFERENCES

[1] J. Malhotra, S. Rani, and P. Sharma, “An intelligent approach for resume ranking using NLP and machine learning,” International Journal of Advanced Computer Science and Applications, vol. 11, no. 5, pp. 458–465, 2020.

[2] S. Gupta and A. Yadav, “AI based resume screening and ranking system,” in Proc. Int. Conf. on Advances in Computing, Communication and Control, 2021, pp. 178–183.

[3] A. K. Sharma and M. Jain, “Automated resume classification using machine learning techniques,” International Journal of Engineering Research & Technology (IJERT), vol. 8, no. 6, pp. 21–25, 2019.

[4] R. K. Singh and P. Verma, “Recruitment using artificial intelligence: A study on automated hiring system,” International Journal of Innovative Technology and Exploring Engineering (IJITEE), vol. 9, no. 3, pp. 420–423, Jan. 2020.

[5] R. Kumar and M. Kumari, “AI-based resume screening system,” in Proc. Int. Conf. on Intelligent Computing and Control Systems (ICICCS), 2022, pp. 345–350.

[6] D. J. Patil and S. Thakur, “Resume shortlisting system using natural language processing,” International Journal of Computer Applications, vol. 183, no. 42, pp. 34–38, 2021.

[7] M. R. Mohan and S. A. Varghese, “Application of NLP in automated resume screening,” in Proc. 6th Int. Conf. on Computing Methodologies and Communication (ICCMC), 2022, pp. 528–532.

[8] N. Tripathi and S. Shukla, “An efficient machine learning model for resume screening,” Journal of Emerging Technologies and Innovative Research (JETIR), vol. 7, no. 9, pp. 102–106, Sep. 2020.

[9] P. Kumar and V. Bansal, “Smart recruitment using AI: A resume parser model,” International Journal of Engineering and Advanced Technology (IJEAT), vol. 8, no. 6S3, pp. 1242–1246, 2019.

[10] S. Deshpande and M. Bendre, “Automated candidate screening using machine learning,” in Proc. Int. Conf. on Advances in Computing, Communication and Control (ICAC3), 2020, pp. 92–96.