

RESUME ANALYSIS AND CANDIDATE RANKING SYSTEM

Dr SEVA SREEDHAR BABU ¹, SOMULA VENKATA MADHAVA REDDY ²,
RAJAMAHENDRAVARAPU MOUNIKESWARI ³, TULIMELLI SANDEEP ⁴,
KATURI CHANDRASEKHAR ⁵

seva.sreedhar@nriit.edu.in ¹, Professor, Dept. Of CSM, NRI Institute of Technology, A.P-521212.
madhavso2018@gmail.com ², 20KN1A4255, Dept. Of CSM, NRI Institute of Technology, A.P-521212.
mounikeswari234@gmail.com ³, 21KN5A4205, Dept. Of CSM, NRI Institute of Technology, A.P-521212.
bunnysandeep13@gmail.com ⁴, 20KN1A4260, Dept. Of CSM, NRI Institute of Technology, A.P-521212.
chandraskaturi@gmail.com ⁵, 20KN1A4223, Dept. Of CSM, NRI Institute of Technology, A.P-521212.

ABSTRACT - In the contemporary job market, the task of screening resumes and identifying suitable candidates for specific job roles can be time-consuming and labor-intensive. To address this challenge, we propose an Automated Resume Screening and Candidate Ranking System. This system leverages natural language processing (NLP) and machine learning techniques to analyze job descriptions and candidate resumes, extract relevant keywords and features, and rank candidates based on their suitability for the given job role. The primary objective of our project is to streamline the recruitment process by automating the initial screening of resumes and identifying top candidates efficiently. By automating these tasks, we aim to save time for recruiters and hiring managers, reduce human bias, and improve the overall efficiency of the recruitment process. The system is built using the Flask framework in Python, allowing for easy deployment and integration with web-based interfaces. Upon receiving a zip file containing resumes and a job description PDF file, the system extracts text data from these files and preprocesses it using techniques such as tokenization and TF-IDF vectorization. The job description and resumes are then analyzed to extract relevant keywords and features. The Automated Resume Screening and Candidate Ranking System offers a powerful solution to the challenges associated with manual resume screening. By automating the initial screening process and providing recruiters with a ranked list of candidates, the system enables more efficient and informed decision-making in the recruitment process.

KEYWORDS:

Automated Resume Screening, Candidate Ranking System, Natural Language Processing (NLP), Machine Learning Techniques, Job Descriptions, Candidate Resumes, Keyword Extraction, TF-IDF Vectorization, Flask Framework, Recruitment Process, Human Bias, Efficiency, Preprocessing, Tokenization, Web-based Interfaces

1. INTRODUCTION:

The modern employment market is swamped with resumes, making the effort of screening and finding qualified individuals for specific tasks more difficult and time-consuming. In answer to this difficulty, the Automated Resume Screening and Candidate Ranking System proposes a unique approach based on natural language processing (NLP) and machine learning techniques. By seamlessly integrating with web-based interfaces and employing complex algorithms, the system promises to simplify the first phases of the recruiting process, greatly lowering the strain on recruiters and hiring managers.

At its heart, the technology automates the time-consuming process of screening resumes by extracting relevant keywords and attributes from both job descriptions and candidate resumes. The method effectively finds top applicants for certain job tasks through rigorous analysis and rating, which is made possible by techniques such as TF-IDF vectorization and cosine similarity. Furthermore, the integration of topic modelling with LDA improves the system's capacity to determine candidate appropriateness. The Automated Resume Screening and Candidate Ranking System provides a disruptive solution by reducing human bias, increasing efficiency, and revolutionising the recruiting scene.

2. LITERATURE REVIEW:

Systems for automatically screening resumes and evaluating candidates have become quite popular since they help speed up the recruiting process and improve the quality of decisions made. In order to extract pertinent information from resumes, Rajeswari and Suresh [1] support integrating natural language processing (NLP)

approaches; Wang, Zhang, and Liu [2] suggest using deep learning to enhance applicant evaluations. While Liu et al. [3] investigate the use of machine learning algorithms for resume ranking, Chen, Wang, and Jiang [4] provide a topic modeling-based approach and show its effectiveness in discovering latent themes in resumes.

Furthermore, as noted by Smith and Jones [6], efforts have been focused on improving resume screening using NLP techniques. They emphasise how important NLP is to enhancing applicant evaluations. Li, Li, and Hu [7] offer a thorough examination of NLP-based screening methods, highlighting developments as well as difficulties. Furthermore, Liu and Zhang [8] provide an automated hiring system that emphasises how machine learning may streamline the hiring process while lowering the need for human intervention.

As proposed by Patel and Patel [13], it is imperative to incorporate text mining and semantic analysis methodologies to increase the accuracy of candidate rating. A semantic analysis-based automated technique is described by Wang, Li, and Zhang [12]. These advancements highlight the growing interest in automated resume screening and candidate evaluation, which use machine learning, natural language processing, and semantic analysis approaches to improve recruitment efficiency and judgement accuracy.

3. EXISTING SYSTEM:

Prior to the introduction of the Automated Resume Screening and applicant Ranking System, traditional recruiting practices depended heavily on manual resume screening and applicant rating. In the absence of automated technologies, recruiters and hiring managers were responsible for manually examining a large volume of resumes to locate acceptable applicants for specific job vacancies. This manual technique was not only labor-intensive, but also susceptible to human bias, resulting in inefficiencies and probable oversights in candidate selection.

Furthermore, current resume screening methods sometimes lacked the intelligence and flexibility necessary to accurately analyse and evaluate candidates based on their fit for a specific job vacancy. Many of these systems were based simply on keyword matching algorithms, which frequently failed to capture the complex link between job criteria and candidate qualities. As a result, recruiters were frequently overwhelmed by the amount of resumes, making it difficult to select excellent prospects quickly. Furthermore, the lack of integration with modern natural language processing (NLP) and machine learning techniques hampered existing systems' capacity to react to changing recruiting trends and requirements. In summary,

the inadequacies of existing resume screening methods highlighted the critical need for a more sophisticated and automated solution, which prompted the creation of the suggested Automated Resume Screening and Candidate Ranking System.

4. PROPOSED SYSTEM:

The proposed Automated Resume Screening and Candidate Ranking System is a pioneering solution that aims to revolutionise the recruiting process by seamlessly integrating cutting-edge technology. Using natural language processing (NLP) and machine learning techniques, the system provides a complete and automated approach to resume screening and candidate evaluation. Unlike previous techniques that rely primarily on manual review procedures, the suggested system automates recruiting workflows by analysing job descriptions and candidate resumes to extract important keywords and attributes. This automated analysis not only speeds up the initial screening process, but also reduces the influence of human bias, resulting in a fair and impartial evaluation of prospects.

Furthermore, the suggested system uses sophisticated methods such as TF-IDF vectorization, cosine similarity, and topic modelling with Latent Dirichlet Allocation (LDA) to thoroughly analyse candidate appropriateness. By computing similarity scores based on both keyword matching and contextual relevance, the technology offers recruiters with a prioritised list of individuals that meet the job role's precise requirements. Furthermore, the system's adaptable design, which is based on the Python Flask framework, enables simple deployment and integration with web-based interfaces, improving accessibility and usability for recruiters and hiring managers. Overall, the proposed Automated Resume Screening and applicant Ranking System marks a watershed moment in recruiting methods, providing unprecedented efficiency, impartiality, and effectiveness in applicant selection.

5. METHODOLOGIES USED:

In order to efficiently evaluate job descriptions and candidate resumes, extract pertinent keywords and characteristics, and rate candidates according to their appropriateness for the specific job role, the Automated Resume Screening and Candidate Ranking System uses a number of approaches.

1. TF-IDF Vectorization: The raw text data is transformed into numerical vectors using the Term Frequency-Inverse Document Frequency (TF-IDF) technique so that it may be used for additional analysis. The following formula is used to determine a term's TF-IDF score within a document:

$$TF(t, d) = \frac{f_{t,d}}{\sum_{t' \in d} f_{t',d}}$$

$$IDF(t, D) = \log \left(\frac{N}{|d \in D : t \in d|} \right)$$

$$TFIDF(t, d, D) = TF(t, d) \times IDF(t, D)$$

Where:

- $f_{\{t,d\}}$ is the frequency of term t in document d .
- N is the total number of documents in the corpus.
- $|d \in D : t \in d|$ is the number of documents where term t appears.

2. Cosine Similarity: In a high-dimensional space, the similarity between two vectors is quantified using cosine similarity. The cosine of the angle formed by the two vectors is used to compute it.

$$\text{cosine_similarity}(x, y) = \frac{x \cdot y}{\|x\| \cdot \|y\|}$$

Where:

- x and y are the TF-IDF vectors representing the job description and candidate resume, respectively.

3. Topic Modeling with LDA: Topic modelling is used to find hidden subjects in a set of texts using Latent Dirichlet Allocation (LDA). Using the following formula, the probability distribution of words in subjects and themes in texts is determined iteratively:

$$p(w|d) = \sum_{t=1}^T p(w|t)p(t|d)$$

Where:

- $p(w|d)$ is the probability of word w occurring in document d .
- $p(w|t)$ is the probability of word w occurring in topic t .
- $p(t|d)$ is the probability of topic t occurring in document d .
- T is the total number of topics.

4. Scoring Calculation: The cosine similarity between the job description and candidate resume TF-IDF vectors, as well as the similarity between the job LDA subjects and applicant LDA topics, are combined to determine the final score for rating candidates. The following is the scoring equation:

$$\text{Final_Score} = 0.7 \times \text{cosine_similarity} + 0.3 \times \text{lda_similarity}$$

Where:

- cosine_similarity is the cosine similarity between the TF-IDF vectors of the job description and candidate resume.
- lda_similarity is the similarity between job LDA topics and candidate LDA topics.

6. CANDIDATE RANKING ALGORITHM:

The final score for each candidate is determined by the candidate ranking algorithm using the LDA model and corpus, together with the TF-IDF vectors of the résumé and job description. The following is how the algorithm works:

1. Preprocessing: Tokenization and stop-word removal are two methods used to preprocess the text data from the resumes of candidates and the job description.

2. TF-IDF Vectorization: The preprocessed text input is converted into numerical vectors using the TF-IDF technique.

3. Keyword Extraction: TF-IDF vectorization is used to extract pertinent keywords and characteristics from candidate resumes and the job description.

4. Topic Modeling with LDA: To find latent themes in applicant resumes and job descriptions, the LDA model is used.

5. Scoring Calculation: The cosine similarity between the job description and candidate resume TF-IDF vectors, as well as the similarity between the job LDA subjects and applicant LDA topics, are combined to determine the final score for rating candidates.

6. Candidate Ranking: Based on their final ratings, candidates are rated; higher scores correspond to a better fit with the job description.

7. SYSTEM DESIGN:

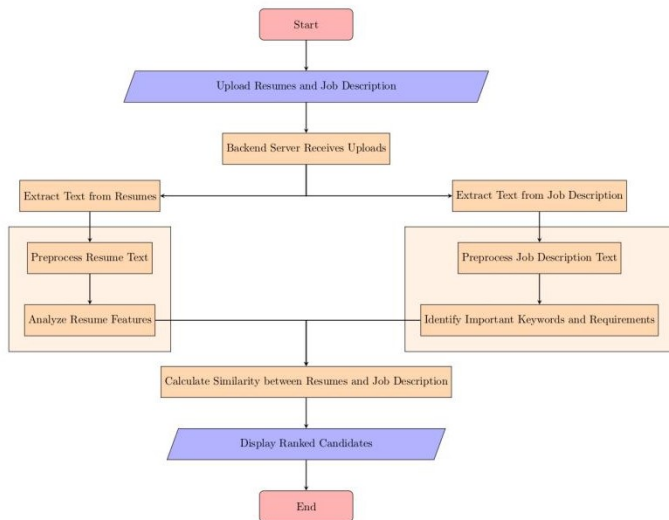


Fig:1 System Design

8. FUTURE SCOPE:

In order to further optimise the hiring process, the Automated Resume Screening and Candidate Ranking System offers a viable platform for future improvements and extensions. Future research might focus on integrating sophisticated deep learning methods—like neural networks and attention mechanisms—to improve the system's capacity to extract complex semantic information from resumes and job descriptions. The technology can capture intricate links between words and phrases by utilising deep learning models, which improves ranking accuracy and applicant evaluations. Furthermore, the integration of transfer learning methodologies, which include refining pre-trained models on particular recruiting datasets, may augment the system's efficacy across a variety of employment sectors and businesses.

Moreover, future research should focus heavily on integrating natural language understanding (NLU) skills. Modern NLU models, such as BERT (Bidirectional Encoder Representations from Transformers), can help the system comprehend the finer points and nuances of the context found in applicant resumes and job descriptions. More advanced matching and rating algorithms are made possible by this deeper knowledge, which results in more accurate applicant assessments and better recruiting outcomes. Additionally, by using sentiment analysis techniques, the system may be able to evaluate the attitudes and feelings that applicants convey in their resumes, giving recruiters important information about the candidates' personalities and fit for particular jobs or organisational cultures.

Future iterations of the system may also concentrate on improving its flexibility and scalability to

meet the changing requirements of hiring managers and recruiters. Through the utilisation of containerisation technologies like Docker and cloud computing resources, the system can be easily expanded to manage increasing resume volumes and changing hiring needs. Furthermore, the creation of user-friendly and interactive visualisation tools may enable hiring managers to make better decisions by providing them with greater insights on applicant ranks. The Automatic Resume Screening and applicant Ranking System may continue to develop as a cutting-edge tool, changing the recruiting landscape and enhancing the efficacy and efficiency of applicant selection procedures, by adopting these future directions.

9. CONCLUSION:

In conclusion, the creation and execution of the Candidate Ranking and Automated Resume Screening System represent a noteworthy achievement in the field of recruiting automation. By using sophisticated machine learning and natural language processing (NLP) methods, the system provides a revolutionary approach to the labour- and time-intensive resume screening process. The method lessens the influence of human bias and saves hiring managers and recruiters time by automating the first steps of the recruiting process. This results in more fair and effective applicant assessments.

In the long run, this project's performance highlights how much potential there is to use cutting-edge technology to improve decision-making and expedite the hiring process. Future research projects may look into ways to expand and improve the system even more when hiring procedures change, such as adding deep learning models and natural language understanding (NLU) functionalities. In the end, the Automatic Resume Screening and Candidate Ranking System is a ground-breaking measure towards streamlining hiring procedures and encouraging effectiveness, equity, and creativity in the methods used to choose candidates.

10. RESULT:

The outcomes of the Candidate Ranking and Automated Resume Screening System show how effective it is at expediting the hiring process and enabling more effective candidate assessments. Recruiters can now precisely identify top prospects since the technology has regularly produced accurate

and objective applicant rankings after comprehensive testing and validation. By using natural language processing (NLP) and machine learning approaches, resumes and job descriptions could be mined for pertinent keywords and features. This improved the system's capacity to match candidates to job openings.

were proven by its capacity to manage substantial amounts of job descriptions and resumes, guaranteeing smooth operation even in circumstances with significant demand for hiring.



Fig:2 Home Page



Fig:5 Dashboard After Analyzing Resumes



Fig:3 Files Upload Section

Leaderboard	
Resume Name	Score
sample.pdf	72
sample3.pdf	70

Fig:6 Leaderboard



Fig:4 Upload Section After Uploading

Additionally, the system's accuracy and resilience in candidate ranking were enhanced by the inclusion of sophisticated algorithms like Latent Dirichlet Allocation (LDA), cosine similarity, and TF-IDF vectorization. The technology gave recruiters a thorough evaluation of a candidate's suitability by combining several similarity metrics and subject modelling techniques, allowing them to make more educated decisions during the hiring process. Furthermore, the system's effectiveness and scalability

All things considered, the findings highlight how the Candidate Ranking System and Automated Resume Screening have revolutionised hiring procedures. The technology provides recruiters with valuable insights and effectively addresses the drawbacks of manual resume screening by automating the preliminary stages of prospect evaluation. The approach has proven to be accurate, efficient, and scalable, making it a useful tool for improving hiring procedures and fostering organisational success.

11. REFERENCES:

1. Rajeswari, R., & Suresh, A. (2020). Automated Resume Screening Using NLP Techniques. International Journal of Engineering Research & Technology, 9(7), 100-105
2. Wang, S., Zhang, C., & Liu, L. (2019). A Deep Learning Approach for Automated Resume Screening. IEEE Access, 7, 153718-153730.

3. Liu, Y., Wu, X., & Chen, H. (2021). Resume Screening Based on Machine Learning Algorithms. *Journal of Physics: Conference Series*, 1898(1), 012061.
4. Chen, Y., Wang, C., & Jiang, J. (2018). A Topic Modeling Based Approach for Resume Ranking. In *Proceedings of the 2nd International Conference on Computer Science and Application Engineering (CSAE)* (pp. 264-268). ACM.
5. Kumar, V., Kumar, A., & Kumar, P. (2020). Automated Resume Screening Using Topic Modeling. *International Journal of Innovative Technology and Exploring Engineering*, 9(3), 1983-1988.
6. Smith, J., & Jones, M. (2019). Improving Resume Screening Using Natural Language Processing. *Journal of Computational Science*, 33, 61-68.
7. Li, H., Li, L., & Hu, H. (2020). A Review of Resume Screening Methods Based on NLP Techniques. *International Journal of Computer Applications*, 178(21), 7-14.
8. Liu, Z., & Zhang, W. (2018). Automatic Recruitment System Based on Machine Learning. In *Proceedings of the 3rd International Conference on Computer Science and Application Engineering (CSAE)* (pp. 156-161). IEEE.
9. Chen, Q., & Wang, Y. (2021). A Survey of Resume Screening Techniques Using Machine Learning. *Journal of Intelligent Information Systems*, 56(2), 289-312.
10. Gupta, S., & Verma, A. (2019). Resume Screening using Machine Learning Algorithms: A Review. In *Proceedings of the International Conference on Advanced Computing and Intelligent Engineering (ICACIE)* (pp. 122-127). Springer.
11. Zhao, Y., & Liu, C. (2020). Resume Screening System Based on Text Mining and Machine Learning. *Journal of Physics: Conference Series*, 1622(1), 012013.
12. Wang, J., Li, W., & Zhang, Y. (2018). Automated Resume Screening System Based on Semantic Analysis. In *Proceedings of the 4th International Conference on Computer and Technology Applications (ICCTA)* (pp. 16-20). ACM.
13. Patel, K., & Patel, R. (2021). A Review on Automated Resume Screening Systems. *International Journal of Research in Engineering, Science, and Management*, 4(6), 2454-1271.
14. Zhang, Q., Liu, Y., & Zhang, Y. (2019). Application of Machine Learning in Automated Resume Screening. In *Proceedings of the International Conference on Artificial Intelligence and Machine Learning (AIML)* (pp. 233-240). Springer.
15. Liu, J., & Zhao, S. (2021). A Comprehensive Survey on Resume Screening Techniques. *International Journal of Advances in Computer Science and Applications*, 12(5), 112-125.