



AI Powered Resume Screening and Candidate Ranking System with Multiple Resume Analysis and PDF Report Generator

¹S Jubeda Banu, ²M Pradeep Kumar Reddy, ³V Radhika, ⁴Y Sunera Tej, ⁵S Mohammed, ⁶S Mohammad Saif

^{1,2,3,4,5,6}Department of Computer Science Engineering (Data Science), GATES Institute of Technology, Gooty, Andhra Pradesh, India

E-mail: ¹jubedashaik123@gmail.com, ²reddypradeepkumar134@gmail.com, ³radhikavennapusala@gmail.com,

⁴sunerataj1013@gmail.com, ⁵sailanimahammed@gmail.com, ⁶shaiksaif123s@gmail.com

Abstract: The AI Powered Resume Screening and Candidate Ranking System is designed to automate the recruitment process by analyzing resumes and comparing them with job requirements. The system extracts important information such as skills, education, and experience to identify suitable candidates. It allows recruiters to upload and process multiple resumes at a time, reducing manual effort and saving time. The system also provides a PDF downloadable option for the results, enabling recruiters to easily review and manage ranked candidate lists.

Keywords: Artificial Intelligence, Resume Screening, Candidate Ranking, Recruitment Automation, Natural Language Processing, Multiple Resume Upload, PDF Downloadable Results.

I. INTRODUCTION

Recruitment is an essential process for organizations to select suitable candidates for job positions. In many companies, recruiters receive a large number of resumes for a single job opening. Manually reviewing these resumes is time-consuming and may lead to errors or difficulties in identifying the most qualified candidates.

With the advancement of technology, Artificial Intelligence (AI) can be used to improve the recruitment process by automating resume screening and candidate evaluation. In this system, AI automates the resume screening process, reducing manual work and improving efficiency. Natural Language Processing (NLP) is used to extract and understand text information from resumes, such as skills, education, and experience. Machine Learning (ML) helps match resumes with job descriptions and ranks candidates based on their relevance to the job role.

The proposed AI Powered Resume Screening and Candidate Ranking System helps recruiters analyze resumes more efficiently. The system allows recruiters to upload and process multiple resumes at a time, which helps in handling large numbers of applications quickly. It extracts key information from resumes and compares it with the job description to identify suitable candidates.

The system ranks candidates based on their matching score and presents the results in an organized format. In addition, the system provides a PDF downloadable option for the results, allowing recruiters to easily download and review the ranked candidate list. This approach improves the efficiency, accuracy, and speed of the recruitment process.

II. RELATED WORK

Several researchers have explored the use of Artificial Intelligence and Natural Language Processing techniques to improve the resume screening and recruitment process. S. Gupta and R. Singh proposed an automated resume screening system using Natural Language Processing to parse resumes and extract relevant features such as skills, education, and experience. Their system ranks candidates according to job descriptions, reducing recruiter bias and improving the efficiency of candidate shortlisting.

Similarly, J. Thomas and A. Mehta developed an AI-driven recruitment system that uses supervised machine learning algorithms such as Random Forest and Support Vector Machine to evaluate candidate profiles. Their approach matches resumes with predefined

competency frameworks and helps recruiters identify suitable candidates more accurately.

P. Sharma and K. Iyer introduced a smart hiring assistant that combines keyword extraction and semantic analysis to evaluate candidate resumes. Their framework ranks candidates based on skill relevance and professional experience, improving the overall hiring process.

Another study by Y. Li and M. Zhang proposed a deep learning approach for automated resume classification. Their model uses neural networks along with embedding techniques such as Word2Vec and BERT to classify resumes into different job categories. This method improves feature representation and significantly reduces errors in candidate shortlisting.

Although these systems improve resume analysis and candidate matching, many existing solutions still lack features for efficiently handling multiple resumes and managing result outputs. The proposed AI Powered Resume Screening and Candidate

Ranking System addresses these limitations by allowing recruiters to upload multiple resumes at a time and providing a PDF downloadable option for the ranked results, making the recruitment process more efficient and user-friendly.

III. PROPOSED SYSTEM

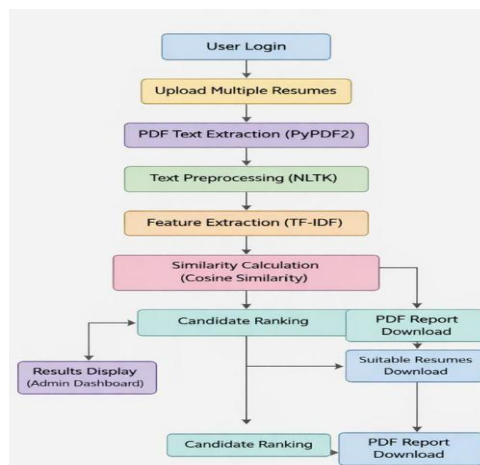
The proposed AI Powered Resume Screening and Candidate Ranking System is designed to automate the recruitment process using Artificial Intelligence (AI), Natural Language Processing (NLP), and Machine Learning (ML) techniques. The system assists recruiters in efficiently analyzing resumes and identifying suitable candidates based on job requirements. The system workflow consists of three major stages: resume parsing, resume–job description matching, and candidate ranking.

3.1 Resume Parsing

In this stage, the system extracts important information from the uploaded resumes. Natural Language Processing techniques are used to identify relevant details such as candidate skills, educational qualifications, work experience, and other important attributes. The system also supports the selection and processing of multiple resumes at a time, which enables efficient handling of large volumes of job applications.

3.2 Resume–Job Description Matching

After extracting the relevant information from the resumes, the system compares the resume content with the job description. Machine Learning techniques are used to calculate the similarity between the candidate profile and the job requirements. This matching process helps determine how well a candidate’s skills and experience align with the required job role.



3.3 Candidate Ranking and Result Generation

Based on the similarity scores obtained from the matching process, the system ranks the candidates according to their relevance to the job description. The ranked results are displayed through the system interface, allowing recruiters to quickly identify the most suitable applicants. Additionally, the system provides a PDF downloadable option, enabling recruiters to download and store the ranked candidate list for further review and documentation.

IV. NLTK

NLTK (Natural Language Toolkit) is used to process the text in resumes and job descriptions. It helps identify important information such as skills, education, and experience. NLTK provides tools like tokenization and stop-word removal to clean and organize the text data. In this system, NLTK is mainly used for text preprocessing, which prepares the resume text for further analysis. By removing unnecessary words and splitting the text into meaningful tokens, the system can focus on important keywords when comparing resumes with job descriptions focus on relevant keywords when comparing resumes with job descriptions.

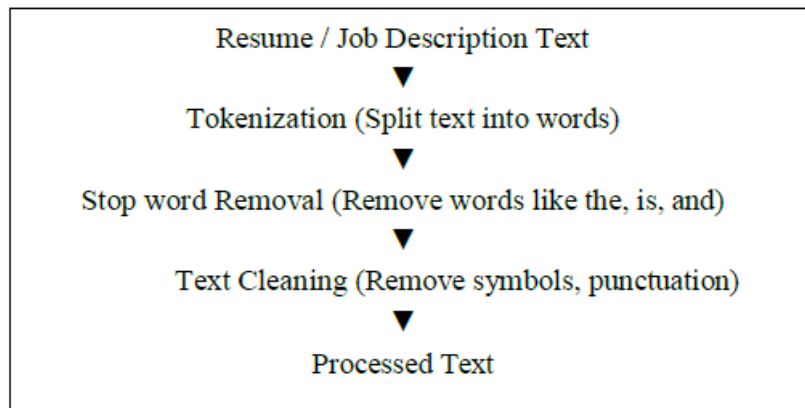


Figure 2: Simple NLTK Text Processing Workflow

TF-IDF (Term Frequency – Inverse Document Frequency)

TF-IDF is a technique used in Natural Language Processing to convert text into numerical values. It helps identify important words in a document by measuring how frequently a word appears in a document compared to how common it is across all documents.

In the resume screening system, TF-IDF is used to convert resume text and job descriptions into numerical vectors. This allows the system to compare resumes with job requirements using mathematical calculations. Words that appear frequently in a specific resume but not in many other resumes receive higher importance. This helps the system focus on important keywords such as skills, tools, and technologies when evaluating candidates.

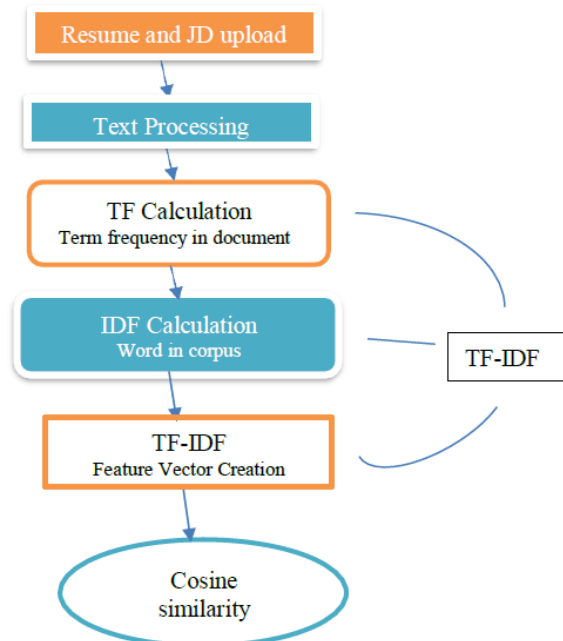


Figure 3: TF-IDF Feature Extraction Process for Resume and Job Description Analysis

COSINE SIMILARITY

Cosine similarity is a widely used technique for measuring the similarity between two text documents. It calculates the cosine of the angle between two vectors in a multidimensional space. In the proposed system, cosine similarity is used to measure the similarity between the TF-IDF vectors of resumes and job descriptions. This helps determine how closely a candidate's resume matches the required job role.

The similarity score ranges from 0 to 1, where a value closer to 1 indicates a higher similarity between the resume and the job description.

The cosine similarity between two vectors (A) and (B) is calculated using the following formula:

$$\text{Cosine Similarity} = \frac{A \cdot B}{|A| |B|}$$

where $(A \cdot B)$ represents the dot product of the vectors, and $(|A|)$ and $(|B|)$ represent the magnitudes of the vectors. Based on the similarity scores obtained, the system ranks candidates according to their relevance to the job description.

PyPDF2

PyPDF2 is a Python library used for reading and extracting text from PDF documents. In the proposed resume screening system, PyPDF2 is used to extract textual information from uploaded resume files in PDF format. Since most resumes are submitted as PDF documents, the library enables the system to automatically read and process resume content without manual intervention.

The extracted text from the PDF resumes is then passed to the Natural Language Processing module for further preprocessing. This process allows the system to identify important information such as candidate skills, education, and work experience. By converting

PDF content into readable text data, PyPDF2 plays an important role in enabling automated resume analysis and candidate evaluation.

PDF RESULT GENERATOR

The PDF Result Generator is used to create a downloadable report of the resume screening results. In the proposed system, the generated results include the ranked list of candidates along with their similarity scores based on the comparison between resumes and job descriptions. This feature allows recruiters to easily download and review the evaluation results for further decision-making.

The system automatically generates a PDF document containing candidate information, ranking order, and matching scores. This helps maintain proper documentation of the recruitment process and improves the usability of the system. The generated PDF report provides a structured and organized summary of candidate evaluation results.

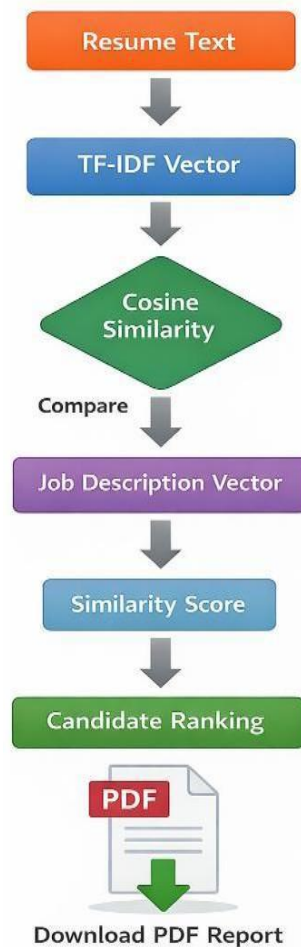


Figure 4: Resume Screening Pipeline Using TF-IDF, Cosine Similarity, and PDF Report Generation

RESULTS

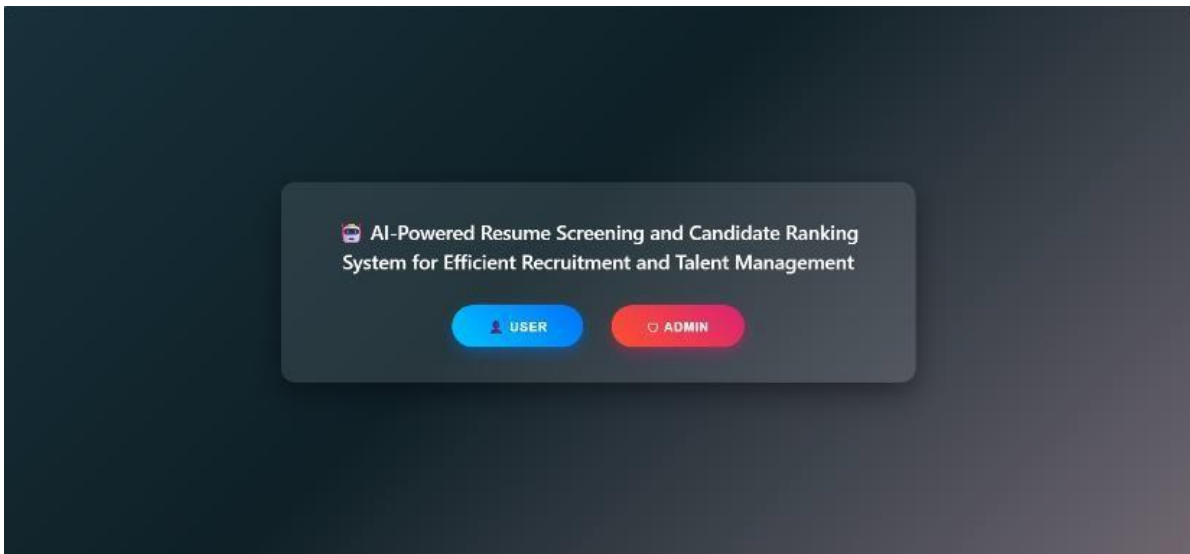
In the proposed project, the results are generated based on the use of Artificial Intelligence (AI), Natural Language Processing (NLP), and Machine Learning (ML) techniques. Artificial Intelligence is used to automate the resume screening process, which reduces manual effort and improves efficiency in recruitment.

Natural Language Processing is used to extract important features from resume text, such as skills, education, and work experience.

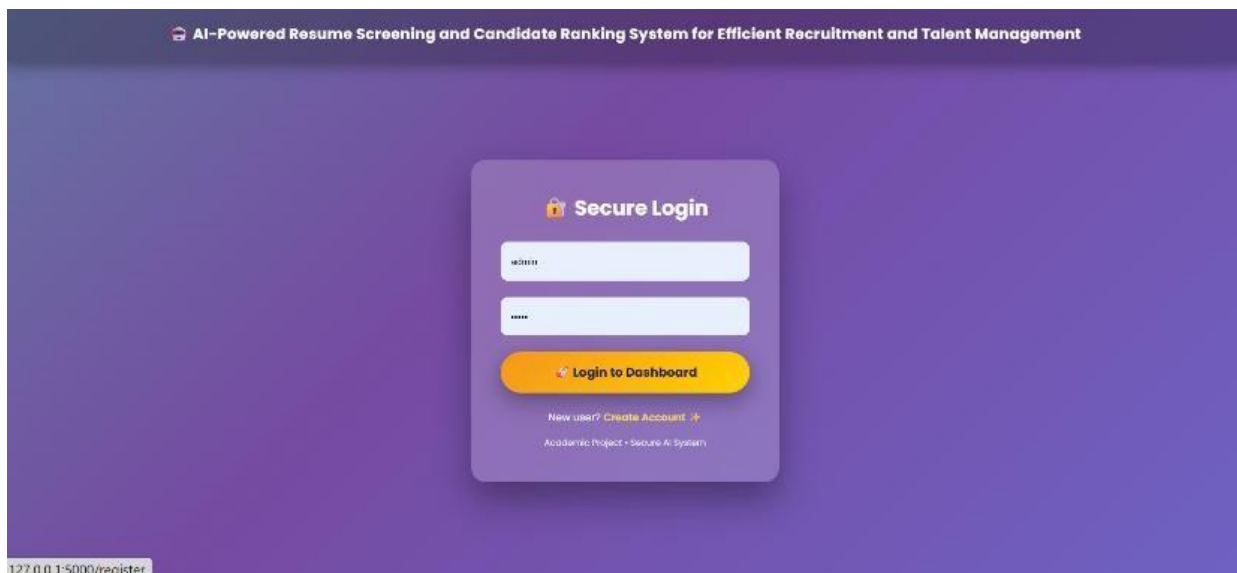
This feature extraction helps the system understand and process the textual information present in resumes and job descriptions.

Machine Learning techniques are used to compare the extracted resume features with the job description and determine the similarity between them. Based on the similarity scores obtained from the matching process, candidates are ranked according to their relevance to the job requirements. The final ranked results are displayed on the dashboard and can also be downloaded as a PDF report.

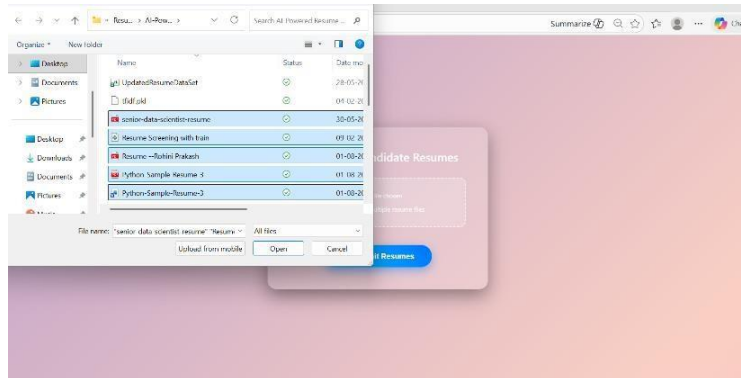
This page shows the login interface of the system. The user can choose either the User or Admin option. The user option is used for uploading resumes, and the admin option is used for checking the screening results.



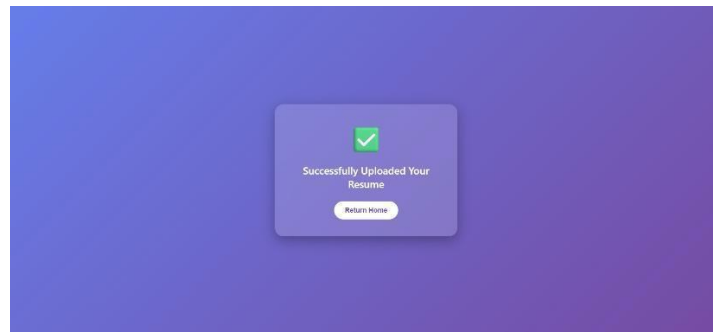
After opening the system, the user must create an account or log in. Once the login is successful, the user can proceed to upload resumes for screening, and the administrator can view the screening results from the dashboard.



The system allows users to select multiple PDF resume files at the same time for screening. After uploading, the system extracts the text from the PDF resumes using the PyPDF2 library. The extracted text is then passed to the Natural Language Processing (NLP) module for preprocessing before performing resume matching and candidate ranking.



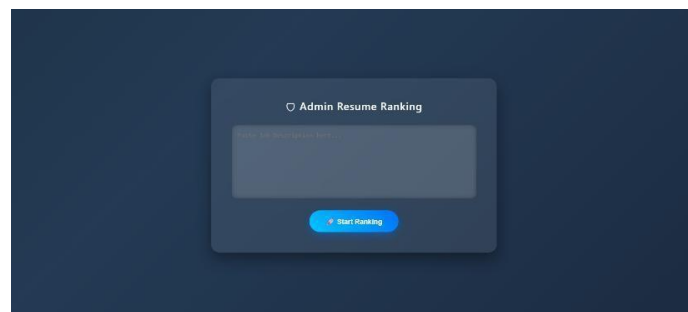
After uploading the resumes, the system shows a success message confirming that the resumes were uploaded successfully. Then the admin can log in and check the results.



This page shows the admin page of the AI-Powered Resume Screening and Candidate Ranking System. After logging in as admin, the administrator enters the job description or job role in the text box.

When the admin clicks the Start Ranking button, the system compares the uploaded resumes with the job description and finds the best matching candidates.

The system uses NLP and ML techniques for resume analysis. NLTK is used for text processing, TF-IDF is used to convert text into vectors, and Cosine Similarity is used to match resumes with the job description. The system then ranks the candidates based on similarity scores.



The results show the rank of each candidate, resume file name, and match score. The candidate with the highest similarity score is ranked first.

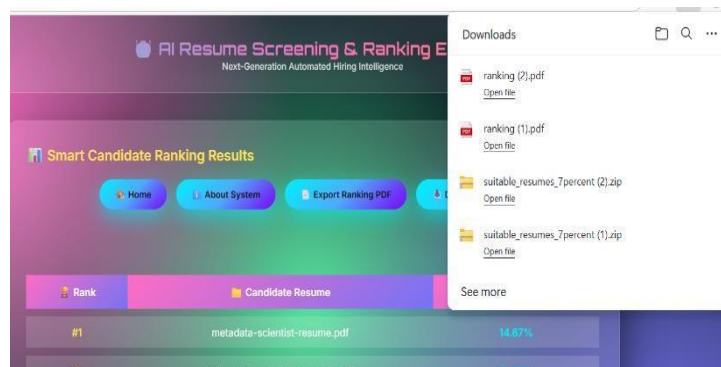
The system uses NLP and Machine Learning techniques to generate these results. The resume text is first processed using NLTK, then TF-IDF converts the text into numerical vectors. The similarity between the resume and the job description is calculated using the

Cosine Similarity algorithm.



This page shows the download options for the resume screening results. After the system ranks the candidates based on similarity scores, the administrator can download the results.

The admin can export the ranking results as a PDF file, which contains the ranked list of candidates and their match scores. The system also allows the admin to download suitable resumes, such as candidates with more than 7% matching score, in a compressed ZIP file. The admin can also download all the resumes if needed.



This feature helps the administrator easily store and review the screening results.

CONCLUSION

In this project, an AI-Powered Resume Screening and Candidate Ranking System was developed to automate the recruitment process. The system allows users to upload multiple resumes and enables administrators to enter job descriptions for candidate evaluation. Natural Language Processing (NLP) techniques are used to extract important information from resumes, while Machine Learning techniques such as TF-IDF and Cosine Similarity are used to compare resumes with job descriptions and calculate similarity scores.

Based on these scores, the system ranks candidates according to their relevance to the job role. The results are displayed on the admin dashboard and can also be downloaded as a PDF report along with suitable resumes. The proposed system helps reduce manual effort, saves time in the recruitment process, and improves the efficiency of candidate selection.

FUTURE WORK

In the future, the proposed system can be improved by integrating advanced machine learning and deep learning techniques to increase the accuracy of resume screening. Additional features such as skill-based filtering, automatic job recommendation, and real-time resume analysis can also be added. The system can be extended to support multiple job roles and larger datasets for better performance. Furthermore, integrating the system with online job portals and recruitment platforms can make the hiring process more



efficient and scalable

REFERENCES

1. Malinowski, S.; Keim, T.; Wendt, O.; Weitzel, T. Matching people and jobs: A bilateral recommendation approach. In Proceedings of the Hawaii International Conference on System Sciences (HICSS), Waikoloa, HI, USA, 4–7 January 2006; pp. 137–146.
2. Paparrizos, J.; Cambazoglu, B.B.; Gionis, A. Machine learned job recommendation. In Proceedings of the ACM Conference on Recommender Systems (RecSys), Barcelona, Spain, 26–30 September 2010; pp. 325–328.
3. Al-Otaibi, A.; Ykhlef, M. Automated resume classification using machine learning techniques. In Proceedings of the IEEE International Conference on Data Science and Advanced Analytics (DSAA), Turin, Italy, 22–25 October 2018; pp. 456–462.
4. Kessler, R.; Brown, P. Resume parsing and job matching using natural language processing techniques. *IEEE Access* 2020, 8, 145231–145240.
5. Zhang, P.; Li, L. Text mining approaches for resume screening in recruitment systems. In Proceedings of the IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 9–12 December 2019; pp. 5874–5879.
6. Jain, S.; Bansal, A. Automated recruitment system using machine learning and NLP techniques. In Proceedings of the IEEE International Conference on Artificial Intelligence and Knowledge Engineering (AIKE), Laguna Hills, CA, USA, 9–11 December 2020; pp. 142–147.
- 7/ Singh, A.; Sharma, R. Intelligent resume ranking system using TF-IDF and cosine similarity. In Proceedings of the IEEE International Conference on Smart Computing and Electronic Enterprise (ICSCEE), Kuala Lumpur, Malaysia, 15–16 July 2021; pp. 1–6.
8. Ahmed, S.; Khan, M. AI-based recruitment system for automated candidate evaluation. *IEEE Access* 2021, 9, 87412–87422.
9. Kumar, S.; Patel, R. Automated resume screening using natural language processing. In Proceedings of the IEEE International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 25–27 March 2022; pp. 1254–1259.
10. Li, X.; Wang, Y. Intelligent hiring system using machine learning and semantic text analysis. In Proceedings of the IEEE International Conference on Data Mining Workshops (ICDMW), Beijing, China, 7–10 December 2019; pp. 532–538.
