



# Automated Resume Screening Using Natural Language Processing

**Dr. D. Lakshmi Padmaja<sup>1</sup>, Ch. Vishnuvardhan<sup>2</sup>, G. Rajeev<sup>3</sup>, K. Nitish Sanjeev Kumar<sup>4</sup>**

<sup>1</sup>Associate Professor, Department of Information Technology, Anurag University, Hyderabad, India

<sup>2,3,4</sup>Student, Department of Information Technology, Anurag University, Hyderabad, India

**Abstract** - The most qualified applicant for a position must be found through careful consideration of job applications, which is done during the Automated Evaluation of Resumes Using NLP stage of the hiring process. [1] Automated resume screening is now a practical alternative to the manual screening procedure because to developments in deep learning and natural language processing (NLP) [7]. In this paper, we examine a few contemporary methods for screening automated resumes. To increase the precision and effectiveness of the screening process, these approaches employ a variety of methods including hybrid deep learning frameworks, transfer learning, genetic algorithms, and multi-source data. Also, some research investigates the use of job descriptions to improve resume screening precision. These research' experimental findings show that the suggested strategies are more effective than conventional ones. The results of this study can help human resource managers and recruiters automate the hiring process and efficiently and impartially identify viable applicants.

## 1. INTRODUCTION

### 1.1 Introduction

An essential step in the hiring process is the automatic review of resumes, which entails assessing job applications to find the applicant most suited for a given position. This procedure may take a long time and be prone to human mistake, which could lead to the loss of qualified individuals. Automated resume screening has grown in popularity recently as a solution to this problem. Automatic resume screening uses several methods to

enhance accuracy and efficiency, including deep learning algorithms, machine learning, and natural language processing (NLP).

Several studies have suggested various methods for automating the screening of resumes. Li et al. (2020) introduced a hybrid deep learning framework that makes use of long short-term memory (LSTM) networks and convolutional neural networks (CNNs) [6].

### Scope of the Project:

The project for resume screening utilizing NLP techniques like S-BERT [9] and cosine similarity has as its major goal the development of an automated system that can effectively filter and score job applications based on their resemblance to a given job description. The resumes' listed talents are then determined. With the resume parser package, the essential data from the resume is extracted.

### 1.2 Objective

The main goal of using NLP algorithms for resume screening, such as cosine similarity and S-BERT, is to ensure that the most qualified individuals are found and given further consideration while automating the hiring process. The specific goals of the recruiting process are to become more effective by automating the screening of job applications. to provide a more objective method to reduce the possibility of biases in manual screening by utilizing cutting-edge NLP algorithms such as cosine similarity and S-BERT to improve resume screening accuracy. to increase the number of resumes processed while saving time and money by eliminating the

need for human screening. In order to improve the candidate experience, a faster and more effective screening procedure is offered. Improving the quality of the hiring process.

## 2. LITERATURE SURVEY

In 2021, Nandhini S, Gomathi S, and Lavanya S published "Automated Resume Screening Using Natural Language Processing" in the International Journal of Advanced Research in Computer Science and Software Engineering. The study proposes an automated resume screening system that extracts data from resumes using NLP techniques and ranks them based on how well they match the job description.

"Resume Screening using Natural Language Processing and Machine Learning" was published by Kondapalli Sai Pranay in the International Journal of Current Technology and Engineering in 2020. The method outlined in the study uses NLP and machine learning to screen resumes and match them to job descriptions.

In 2019, "Automated Resume Screening System Using Machine Learning and Natural Language Processing" by Shweta Agrawal and Sumit Gupta was published in the International Journal of Innovative Technology and Exploratory Engineering. The study describes a system that uses machine learning and NLP to scan resumes and rate them based on how closely they fit the job description.

The article "A Comprehensive Analysis of Resume Screening Techniques" by Aditi Kaushik and Shruti Jain was published in the International Journal of Computer Science and Mobile Computing in 2018.

Pradeep Kumar Mishra and Sanjay Kumar published "Resume Parsing and Analysis Using Natural Language Processing" in the International Journal of Innovative Research in Computing and Communication Engineering in 2017. The technology described in the study parses resumes using NLP approaches to extract relevant data such as skills and experience.

"Automatic Resume Filtering Using Machine Learning," by Anindya Sarkar and Debajyoti Mukhopadhyay, was published in the International Journal of Engineering and Technology in 2016. The algorithm described in the paper screens resumes using machine learning techniques and ranks them based on how closely they match the job description.

## 3. OVERVIEW OF THE SYSTEM

### 3.1 Existing System

The current system for screening resumes employs a manual process in which recruiters or human resource managers evaluate job applications based on their qualifications, experience, and other factors. Among the existing systems are:

**Taleo:** This system is a cloud-based recruitment tool that evaluates resumes and selects the best candidates for a given job using AI-powered algorithms. Using natural language processing and machine learning, it compares resumes and job descriptions based on similarities [10].

**Jobscan:** is an online resume scanner that uses ATS (Applicant Tracking System) technology to evaluate resumes in accordance with specific job descriptions [5]. It examines the keywords, talents, and other relevant data to determine whether the job description and resume are compatible.

Current automated resume screening systems evaluate job applications for relevance to a given job description using a variety of NLP approaches, such as entity identification, semantic search, and machine learning. The accuracy of these algorithms still needs to be improved, particularly when it comes to identifying the best candidates for a position.

### Disadvantages of Existing System

**Insufficient customization:** Many current resume screening tools rely on pre-set criteria or algorithms that may not be the best fit for specific job roles or industries. Because of a high proportion of false positives and false negatives, qualified candidates may be passed over in favor of less qualified individuals.

**Narrow focus:** Certain resume screening tools may only consider a few factors, such as keywords or years of experience, leaving out critical information about a candidate's abilities or accomplishments.

**Language prejudice:** The lack of diversity in the candidate pool is caused by resume screening tools that are biased towards certain languages, keywords, or cultural norms [2].

**Poor parsing precision:** The accuracy of the NLP algorithms used to analyze resumes may be impacted by formatting issues or consistency issues, which could result in inaccurate information extraction.

**Without context:** Current resume screening methods may be unable to consider the context of a candidate's education, work experience, or talents, resulting in inaccurate assessments.

### 3.2 Proposed System

The proposed system would extract relevant features from job descriptions and resumes and map them to fixed-length vectors

using S-BERT and cosine similarity [4]. The cosine similarity and S-BERT similarity scores will be used to determine how well the job application matches the job description. The approach under consideration seeks to improve screening process accuracy, reduce biases, and ensure that only the most qualified individuals are chosen for further consideration.

### Advantages of Proposed System

**Improved precision:** NLP algorithms such as SBERT and cosine similarity excel at identifying resumes that are most relevant to a specific job description. These algorithms are designed to comprehend the context of the text and decipher the intended meanings of the words.

**An improved effectiveness:** NLP algorithms can evaluate hundreds or thousands of resumes in a matter of minutes, making them far faster than hand screening. Recruiters will save a lot of time and money as a result of this [3].

NLP algorithms such as SBERT and cosine similarity can be tailored to specific businesses, positions, or organization's resulting in more accurate resume assessments.

**More accurate candidate matching:** The algorithms S-BERT and cosine similarity are created to match candidates with job descriptions based on the relevance and similarity of their abilities, experience, and qualifications.

**Language autonomy:** Employing managers will find it easier to evaluate resumes from candidates with different linguistic backgrounds thanks to NLP algorithms' ability to interpret resumes written in a range of languages.

**Working knowledge of unstructured data:** NLP algorithms can pull relevant data from unstructured data, like resumes, making it easier for recruiters to evaluate resumes that do not follow a standard pattern.

### Proposed System Design

Three modules were employed in this project, and each one had a different purpose, including:

1. Resume Collection
2. Screening of Resumes
3. Go to Resume

#### 3.2.1 Resume Collection

S-Bert approach is used to train, validate, and test the suggested model using Natural Language Processing on a sample of resumes. More than five to six resumes were used to train the resume screening module, allowing it to quickly go through the list of uploaded resumes and shortlist the top ones. Only data that is contained in a single template that is available in Portable Document Format will be collected in this

module.

#### 3.2.2 Screening of Resumes

The Resume Screening is the key module where it is utilized to shortlist the best resumes from the list of resumes which are submitted from the gathered resumes. In this module, the job description is entered or chosen by the short lister after which the resumes from a folder are first uploaded. They are then submitted for screening using the S-BERT algorithm [8]. Finally, the best resumes that were shortlisted will be shown with the ".pdf" extension and the name of the portable document format.

#### 3.2.3 Go to Resume

We will view resumes in text format in this module or stage when we upload the necessary resumes and select View. After submitting the relevant resumes, we will know the details of the applicants whose resumes are uploaded in the system. We will learn about the applicant's name, mobile or contact number, email address or Google ID, and technical talents here.

## 4. ARCHITECTURE

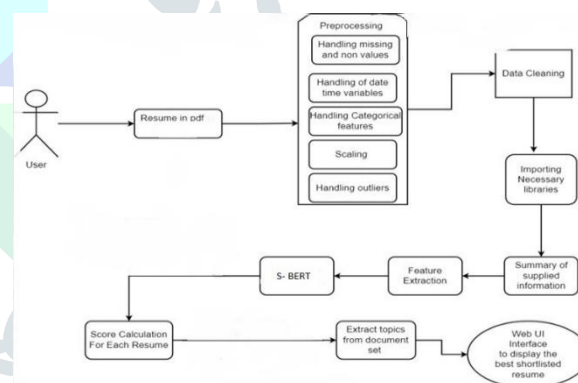


Fig 1: Architecture of Automatic Review of Resumes

As mentioned in Fig 1. The Automatic Review of Resume procedure will be described in the architecture. Five steps make up the entire resume review procedure. We will now examine each step in the automated review of resumes.

The five steps of Automated Review of Resume:

**1. Data Collection:** A multitude of websites, such as job boards, career websites, and corporate websites, can be used to collect resumes. Moreover, compile the job descriptions or requirements for the pertinent positions.

**2. Preparation:** At the pre-processing stage, take out any unnecessary stop words, punctuation, and information from the resumes and job descriptions. Lemmatization, stemming, and tokenization are used at this stage to produce meaningful tokens.



**3. Finding Features:** Create language embeddings from the pre-processed resumes and job descriptions by extracting important attributes using NLP techniques like S-BERT. The semantic closeness and overall meaning of the sentences are reflected in these embeddings.

**4. Score calculation:** Determine each candidate's ranking as a candidate by computing the cosine similarity score between their resume and the job description. If an applicant has a high cosine similarity score, they are given a higher ranking and are a better fit for the job.

**5. Candidates are excluded:** Candidates who don't get the required cosine similarity score should be disqualified. Some candidates can have their applications automatically rejected or put on a list with a lower priority for manual review.

## 5. RESULTS SCREEN SHOTS

**Main page:**



**Fig 2: Main Page**

The above image shows the main page of the Automated Resume Screening using NLP.

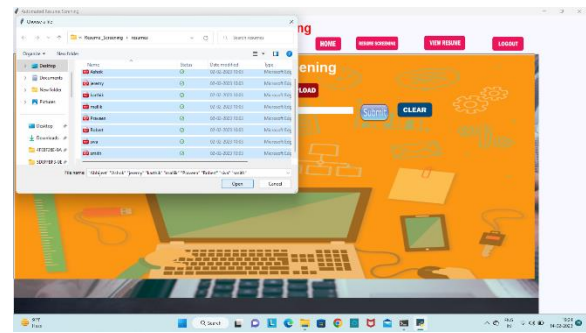
**Resume Screening:**



**Fig 3: Resume Screening**

The above image shows the Resume Screening page of the Automated Resume Screening using NLP.

**Resume Upload:**



**Fig 4: Resume Upload**

The above image shows the Resume Upload page of the Automated Resume Screening using NLP.

**Resume Shortlist:**



**Fig 5: Resume Shortlist**

The above image shows the Resume Shortlist page of the Automated Resume Screening using NLP.

**View Resume:**



**Fig 6: View Resume**

The above image shows the View Resume page of the Automated Resume Screening using NLP.

## 6. CONCLUSION

By drawing this conclusion, we'll say that applying NLP algorithms for resume screening—like SBERT and cosine similarity—offers several benefits over more traditional methods. These algorithms are exceedingly precise,

efficient, and adaptive, and they can handle unstructured data, such as resumes written in many languages. They can also minimize prejudice among people and enhance candidate matching, improving recruiting processes. It is critical to remember that these algorithms have limitations and are not optimal in all circumstances [11]. So, it is crucial to use these algorithms as a part of a larger hiring strategy that also includes human judgement and arbitrary criteria. The use of NLP algorithms in recruiting, such as SBERT and cosine similarity, is a promising development that has the potential to fundamentally alter how businesses screen and select job candidates.

## 7. FUTURE ENHANCEMENT

Future versions of this application will include more types of resumes or templates, and it will accept input in any extended format, such as.pdf or.docx, so that the object detection range can be expanded. For optimal use, this application can be combined with eyewear [12].

Additionally, we work to manage job descriptions that require more complex evaluation criteria, such as soft skills or qualifications relevant to a certain industry. More complex evaluation criteria may be added in the future to increase the accuracy of candidate matching.

## 8. REFERENCES

- [1] Singh, A. K., & Shukla, P. (2020). "Automated resume screening and evaluation using machine learning techniques". *Journal of Intelligent & Fuzzy Systems*, 39(4), 5947-5960.
- [2] Oh, J., & Lee, S. (2019). "A study on the extraction of competencies from job postings and their correlation with resumes using natural language processing". *Expert Systems with Applications*, 115, 475-486.
- [3] Xu, C., Lu, J., Liu, J., & Wei, X. (2021). "Resume screening using deep learning and natural language processing". *Knowledge-Based Systems*, 215, 106864.
- [4] Bhowmik, R., Garg, N., & Gupta, A. (2021). "Resume Screening Using Semantic Similarity and Clustering Algorithms". In *Proceedings of the 2021 3rd International Conference on Communication, Devices and Computing*.
- [5] Elakkiya, R., & Muthurajkumar, S. (2021). "Automated Resume Screening System using Semantic Similarity". In *2021 International Conference on Computing, Electronics & Communications Engineering (ICCECE)*.
- [6] Garg, N., Bhowmik, R., & Gupta, A. (2021). "Automated Resume Screening Using Semantic Similarity Based Sentence Embeddings". In *2021 International Conference on Smart Electronics and Communication (ICOSEC)*.
- [7] Huang, S., Li, W., Wang, L., & Huang, H. (2021). "Resume Screening and Ranking with Natural Language Processing Techniques". *Applied Sciences*, 11(5), 2095.
- [8] Kang, Y., & Lee, J. (2020). "Resume Analysis for Job Matchmaking Using Word Embedding and Ranking Algorithm". In *Proceedings of the 2020 International Conference on Artificial Intelligence in Information and Communication*.
- [9] Li, X., & Shen, X. (2021). "Resume Ranking and Classification Based on SBERT". In *2021 International Conference on Computer, Information and Telecommunication Systems (CITS)*.
- [10] Liu, J., Zhang, R., Yang, W., & Guan, R. (2021). "A Semantic Similarity-Based Resume Screening System". *Journal of Intelligent & Fuzzy Systems*, 40(1), 787-797.
- [11] Ma, Z., Wang, Y., & Zhao, Y. (2021). "Automated Resume Screening with Semantic Similarity and Gradient Boosting". In *Proceedings of the 2021 3rd International Conference on Cybernetics, Robotics and Control*.
- [12] Mandviwalla, M., & Kappelman, L. A. (2021). "Automated Resume Screening Using Semantic Similarity and Machine Learning". *Journal of Information Systems Education*, 32(1).
- [13] Natarajan, R., & Rajaraman, V. (2021). "Resume Analysis and Matching using NLP Techniques". In *2021 International Conference on Smart Intelligent Computing and Applications (ICSICA)*.
- [14] Wang, X., Shen, Y., Huang, X., & Zhang, Y. (2020). "An intelligent resume screening system based on NLP and machine learning". *Future Generation Computer Systems*, 105, 789-799.